

Europäisches Patentamt

European Patent Office

Office européen des brevets



(11) **EP 1 024 200 A2**

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
02.08.2000 Bulletin 2000/31

(51) Int. Cl.⁷: **C12Q 1/68**

(21) Application number: **00250023.9**

(22) Date of filing: **26.01.2000**

(84) Designated Contracting States:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE**
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **27.01.1999 US 238402**

(71) Applicant:
**Affymetrix, Inc. (a California Corporation)
Santa Clara, CA 95051 (US)**

(72) Inventors:
• **Patil, Nila**
Woodside, CA 94062 (US)
• **Shah, Nila**
Cupertino, CA 95014 (US)
• **Warrington, Janet A.**
Los Altos, CA 94024 (US)

(74) Representative:
UEXKÜLL & STOLBERG
Patentanwälte
Beselerstrasse 4
22607 Hamburg (DE)

(54) **Genetic compositions and methods**

(57) This invention provides nucleic acid segments derived from the human genome, including polymorphic sites. Allele-specific primers and probes hybridizing to these sites and to regions flanking these sites are also provided. This invention further provides methods of analyzing a nucleic acid from an individual or a group of individuals. The nucleic acids, primers, and probes are useful tools for applications including forensics, paternity testing, medicine, e.g., the correlation of polymorphisms with phenotypic traits, and genetic analysis, e.g., genetic mapping of such phenotypic traits.

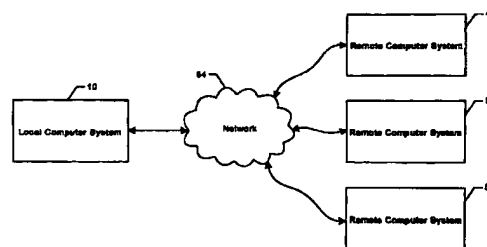


Fig. 1B

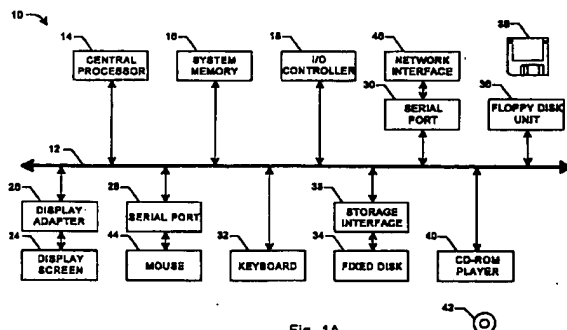


Fig. 1A

EP 1 024 200 A2

Description**CROSS-REFERENCES TO RELATED APPLICATIONS**

- 5 [0001] The present application claims priority from USSN 09/238,402 filed January 27, 1999, the disclosure of which is incorporated by reference.

BACKGROUND OF THE INVENTION

- 10 [0002] The genomes of all organisms undergo spontaneous mutation in the course of their continuing evolution generating variant forms of progenitor sequences (See, e.g., Gusella, *Ann. Rev. Biochem.* 55, 831-854 (1986)). A variant form may be neutral or confer an evolutionary advantage or disadvantage relative to a progenitor sequence. In some instances, a variant form confers a lethal disadvantage. In other instances, a variant form confers an evolutionary advantage to the species and is eventually incorporated into the deoxyribonucleic acid (DNA) of many or most members of the species and effectively becomes the progenitor sequence. In many instances, both the progenitor and the variant form(s) survive and co-exist in a species population. The coexistence of multiple forms of a sequence gives rise to polymorphisms.

- [0003] Several different types of polymorphism have been reported. A restriction fragment length polymorphism (RFLP) is a variation in a DNA sequence that alters the length of a restriction fragment (see, e.g., Botstein et al., *Am. J. Hum. Genet.* 32, 314-331 (1980)). The RFLP may create or delete a restriction site, thus changing the length of the restriction fragment. RFLPs have been widely used in human and animal genetic analyses (See, e.g., WO 90/13668; WO 90/11369; Donis-Keller, *Cell* 51, 319-337 (1987); and, Lander et al., *Genetics* 121, 85-99 (1989)). Where a heritable trait can be linked to a particular RFLP, the presence of such RFLP in a human can be used to predict the likelihood that the individual will also exhibit the trait.

- 25 [0004] Other polymorphisms take the form of short tandem repeats (STRs), for example, tandem di-, tri- and tetra-nucleotide repeated motifs. These tandem repeats are also referred to as variable number tandem repeat (VNTR) polymorphisms. VNTRs have been used in identity and paternity analyses as well as in a large number of genetic mapping studies. (See, e.g., US 5,075,217; Armour et al., *FEBS Lett.* 307, 113-115 (1992); Horn et al., WO 91/14003; Jeffreys, EP 370,719).

- 30 [0005] Other polymorphisms take the form of single nucleotide variations between individuals of the same species. Such polymorphisms are far more frequent than RFLPs, STRs and VNTRs. Some single nucleotide polymorphisms (SNPs) occur in protein-coding sequences, in which case, one of the polymorphic forms may give rise to the expression of a defective, or other variant, protein and, potentially, a genetic disease. Examples of genes, in which polymorphisms within coding sequences give rise to genetic disease include α -globin (sickle cell anemia) and CFTR (cystic fibrosis). 35 SNPs also occur in noncoding regions and these SNPs may result in defective protein expression (e.g., as a result of defective splicing). Further, some SNPs have no phenotypic effects.

- [0006] SNPs can be used in the same manner as RFLPs and VNTRs but offer several advantages. SNPs occur with greater frequency and are spaced more uniformly throughout the genome than other polymorphisms. The greater frequency and uniformity of SNPs mean that there is a greater probability that such a SNP will be found in close proximity to a genetic locus of interest than would be the case for other polymorphisms. Also, the different forms of characterized SNPs are often easier to distinguish than other polymorphisms (e.g., by use of assays employing allele-specific hybridization probes or primers).

- 45 [0007] Despite the increased amount of nucleotide sequence data being generated in recent years, only a minute proportion of the total repository of polymorphisms in humans and other organisms has been identified. The paucity of polymorphisms hitherto identified is due to, in substantial part, the large amount of work required for their detection by conventional methods. A conventional approach to identifying polymorphisms, for example, sequences the same stretch of oligonucleotides in a population of individuals by dideoxy sequencing. In this type of approach, the amount of work increases in proportion to both the length of the sequence and the number of individuals in a population and, as those skilled in the art will understand, becomes impractical for large stretches of DNA or large numbers of subjects.

- 50 [0008] All documents, i.e., publications and patent applications, cited in this disclosure, including the foregoing, are incorporated by reference herein in their entireties for all purposes to the same extent as if each of the individual documents were specifically and individually indicated to be so incorporated by reference herein in its entirety.

SUMMARY OF THE INVENTION

- 55 [0009] This invention provides nucleic acid segments of between 10 and 100 contiguous bases from a fragment shown in TABLE 1, column 1 including a polymorphic site or an immediately adjacent base. Complements of these segments are also provided in this invention. The segments can be DNA or ribonucleic acid (RNA), and can be double- or

single-stranded. Some segments are 10-20 or 10-50 bases length. Preferred segments include a diallelic polymorphic site. The base occupying the polymorphic site in the segments can be the reference (TABLE 1, column 3) or an alternative base (TABLE 1, column 5).

[0010] The invention further provides allele-specific oligonucleotides that hybridizes to a segment of a fragment shown in TABLE 1, column 8 or its complement. These oligonucleotides can be probes or primers. Also provided are isolated nucleic acids comprising a sequence shown in TABLE 1, column 8, or the complement thereto, in which the polymorphic site within the sequence is occupied by a base other than the reference base shown in TABLE 1, column 3.

[0011] The invention further provides methods of analyzing a nucleic acid from a subject, which may be, but is not limited to a human being. The novel methods determine which base is present at any one of the polymorphic sites shown in TABLE 1. Optionally, a set of bases occupying a set of the polymorphic sites shown in TABLE 1 is determined. These types of analyses can be performed on a plurality of subjects who are tested for the presence of a disease phenotype. The presence or absence of disease phenotype can then be correlated with a base or set of bases present at the polymorphic sites in the subjects tested.

DEFINITIONS

[0012] An oligonucleotide, as the case may be, can be DNA or RNA, and single-or double-stranded. Oligonucleotides can be composed of naturally occurring or synthetic bases, but in either case are generally prepared by synthetic means. Preferred oligonucleotides of the invention include segments of DNA, or their complements including any one of the polymorphic sites shown in TABLE 1. The segments are usually between 5 and 100 bases in length, and often between 5-10, 5-20, 10-20, 10-50, 20-50 or 20-100 bases in length. The polymorphic site can occur within any position of the segment. The segments can be from any of the allelic forms of DNA shown in TABLE 1.

[0013] Hybridization probes are oligonucleotides capable of binding in a base-specific manner to a complementary strand of nucleic acid. Such probes include peptide nucleic acids. (See, e.g., Nielsen et al., *Science* 254, 1497-1500 (1991)).

[0014] A primer is a single-stranded oligonucleotide capable of acting as a point of initiation for template-directed DNA synthesis under suitable conditions e.g., buffer and temperature, in the presence of four different nucleoside triphosphates and an agent for polymerization, such as, for example, DNA or RNA polymerase or reverse transcriptase. The length of the primer, in any given case, depends on, for example, the intended use of the primer, and generally ranges from 15 to 30 nucleotides. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with the template. A primer need not reflect the exact sequence of the template but must be sufficiently complementary to hybridize with such template. The primer site is the area of the template to which a primer hybridizes. The primer pair is a set of primers including a 5' upstream primer that hybridizes with the 5' end of the sequence to be amplified and a 3' downstream primer that hybridizes with the complement of the 3' end of the sequence to be amplified.

[0015] Linkage describes the tendency of genes, alleles, loci or genetic markers to be inherited together as a result of their location on the same chromosome, and can be measured by percent recombination between the two genes, alleles, loci or genetic markers.

[0016] Polymorphism refers to the occurrence of two or more genetically determined alternative sequences or alleles in a population. A polymorphic marker or site is the locus at which divergence occurs. Preferred markers have at least two alleles, each occurring at frequency of greater than about 1 percent (%), and more preferably greater than about 10% or 20% of a selected population. A polymorphic locus may be as small as one base pair. Polymorphic markers include RFLPs, variable number of tandem repeats VNTRs, hypervariable regions, minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats, simple sequence repeats, and insertion elements such as, for example, Alu. The first identified allelic form is arbitrarily designated the reference form while subsequently identified allelic forms are designated as alternative or variant alleles. The allelic form occurring most frequently in a selected population is often referred to as the wildtype form. Diploid organisms may be homozygous or heterozygous for allelic forms. A diallelic polymorphism has two forms. A triallelic polymorphism has three forms.

[0017] A SNP occurs at a polymorphic site occupied by a single nucleotide, which is the site of variation between allelic sequences. This site of variation is usually both preceded by and followed by highly conserved sequences e.g., sequences that vary in less than 1/100 or 1/1000 members of the populations of the given allele.

[0018] A SNP usually arises due to the substitution of one nucleotide for another at the polymorphic site. These substitutions include both transitions (i.e. the replacement of one purine by another purine or one pyrimidine by another pyrimidine) and transversions (i.e. the replacement of a purine by a pyrimidine or vice versa). SNPs can also arise from either a deletion of a nucleotide or from an insertion of a nucleotide relative to a reference allele.

[0019] Hybridizations, e.g., allele-specific probe hybridizations, are generally performed under stringent conditions. For example, conditions where the salt concentration is no more than about 1 Molar (M) and a temperature of at least 25 degrees-Celsius (°C), e.g., 750 mM NaCl, 50 mM NaPhosphate, 5 mM EDTA, pH 7.4 (5X SSPE) and a temperature

of from about 25 to about 30°C.

[0020] An isolated nucleic acid is an object species invention that is the predominant species present (i.e., on a molar basis it is more abundant than any other individual species in the composition). Preferably, an isolated nucleic acid comprises at least about 50, 80 or 90% (on a molar basis) of all macromolecular species present. Most preferably, the object species is purified to essential homogeneity (contaminant species cannot be detected in the composition by conventional detection methods).

DESCRIPTION OF THIS INVENTION

I. NOVEL POLYMORPHISMS OF THIS INVENTION

[0021] The novel polymorphisms of this invention are listed in TABLE 1. The second column of TABLE 1 lists the names assigned to the fragments in which the polymorphism occurs. Sequences of fragments beginning with "HT" are available from the TIGR database which is accessible at:

http://www.tigr.org/tdb/hgi/searching/hgi_reports.html, and those starting with "STSG" are from the Whitehead database which is accessible at http://carbon.wi.mit.edu:8000/cgi-bin/contig/sts_info?database=release. Sequences can also be obtained from the GenBank database accession number which may be obtained by using the BLAST program available at <http://www.ncbi.nlm.nih.gov/BLAST>. The fragments are all human genomic fragments. Also included within the scope of this invention are all analogous fragments of other species to such human genomic fragments of TABLE 1.

[0022] The second column of TABLE 1 lists the position of a polymorphic site within the fragment. These positions are numbered consecutively with the first base of the fragment sequence assigned the number one. The third column of TABLE 1 lists the base occupying the polymorphic site in the sequence in the data base. This base is arbitrarily designated the reference, or prototypical form, but is not necessarily the most frequently occurring form. The fifth column of TABLE 1 lists other base(s) found at the polymorphic site. The eighth column of TABLE 1 lists about 15 bases of sequence on either side of the polymorphic site in the fragment. These sequences are designated successive SEQ ID Nos. starting at SEQ ID No. 1. The sequences can be either DNA or RNA. In the latter, the T's shown in TABLE 1 are replaced by U's. The base occupying the polymorphic site is indicated using IUPAC-IUB nomenclature.

FragmentCode	Position	RefAllele	FreqP	AltAllele	FreqQ	Heterozygosity	SequenceTag	ReversePrimer
D42043	1470	G	0.92	A	0.08	0.15	ACCAGCGAGGGGAGTRTATCCACCAAGCAGA	
D42043	1697	A	0.54	G	0.46	0.5	GGAAACTGTGTTCRGGACAGCAGCAGGAG	
D42043	2303	A	0.81	T	0.19	0.31	GAGATTCTGGGTCAA WTCCATGGACACCCCTG	
D45248	523	C	0.95	G	0.05	0.09	GAGGTGAATGCCGTSAA GACCAAAAGTGGAA	
M69066	2504	C	0.72	G	0.28	0.4	ATGGTGCTTCAAGASTTCAACCAAAATAC	
M69066	2874	A	0.78	G	0.22	0.35	CTCCACAGCCATCCCRGCCCTTGGCATCTAGA	
M69066	3321	A	0.78	C	0.22	0.35	GATTTGGCCTCATMCTTTACCAACCCAC	
U56637	2228	G	0.13	A	0.88	0.22	AGCAAGTTGTTCARGTGACCCATTGAGCT	
X12671	18	T	0.39	A	0.61	0.48	AGGCGAAGGTAGGCWGGCAGATACGTTCTGT	
X12671	1504	A	0.94	T	0.06	0.1	CAGGTTCTATTTGGAWTTTATATACAACTGT	
AB006782	93	G	0.82	A	0.18	0.29	GAGATGGCCTTCAGCRGTTCCAGGCTCCCT	
AB006782	1622	T	0.86	A	0.14	0.24	AGCAGGCTGTATGGCWTCCTGCTGGCTCCA	
AB006782	1641	C	0.74	A	0.26	0.39	CACCTGGCTCCACCCAMCTGACCAAGTGTTC	
AB006782	1658	C	0.91	T	0.09	0.16	TGACCAAGAGTGTTCCTTTCAGAGGACTGGCT	
AB006782	1664	A	0.62	G	0.38	0.47	GAGTGTCTCTTCAGRGGACTGGCTCCTTTC	
HT3708	1303	C	0.94	G	0.06	0.1	GGAAACCTCAJTGTTSTGTACAAAGTACTAG	
HT3708	1645	C	0.53	T	0.47	0.5	TGAAACTCATCTCATYACCTTTTCTGGGT	
J02621	368	C	0.71	T	0.29	0.41	AGAACTAAAGAAAGAYTTACCTGGGAAAAAC	
J02621	377	G	0.5	A	0.5	0.5	AGAAAGACTTACCTGCRGAAAAACGGGAAACG	
J02621	522	C	0.79	G	0.21	0.34	GAGGAAATTTTATTSAACTATTTTGTAAATG	
J02621	530	T	0.86	C	0.14	0.24	TTTTATCAACTATTYTGTAATGCAAGTTTT	
J02621	659	T	0.38	C	0.63	0.47	TTGTTTATTTTGGYACAAACCAAGAAATAG	
J02621	669	A	0.91	G	0.09	0.16	TTTGGTACAAACCAAGAAATAGTGTGGGATAT	
J02621	674	G	0.86	C	0.14	0.24	TACAAACCAAGAAATASGTGGGATATTGAAAT	
J02621	691	A	0.95	G	0.05	0.1	GTGGGATATTGAAATRTGGGAGGCTCTGACT	
J02621	701	C	0.85	T	0.15	0.25	GAATTA TGGGAGGCTYTGA CTGTCTCGGGTG	
J02621	777	A	0.91	C	0.09	0.16	AATACAAAGCATATTMAATGGCAATATGGAG	
J02621	838	A	0.64	C	0.36	0.46	TAAATTACTTCTATTMCCATGTGTGTTTTTA	
J02621	842	T	0.85	C	0.15	0.25	TTACTTCTATTACCA YGTTGTTTTTTTAGTAG	
J02621	851	T	0.91	G	0.09	0.16	TTACCATGTGTTTITKTAGTAGAAATTGTTTC	
J02621	921	T	0.91	G	0.09	0.16	GAA TTGTGTGCACCTCKGTAAACATCTTTGGTT	
J02621	1014	A	0.42	G	0.58	0.49	TGTAGTGTACGTGCTRTTGA GTTGTGA ACTG	
K01160	45	T	0.56	C	0.44	0.49	GATGCTGGGGGGCCCTYGCCCTGACCAACCGTG	
K01160	62	T	0.91	C	0.09	0.17	CCCTGACCAACCGTGAYGAGCCCTGTGTGGAGG	

K01160	69	C	0.7	T	0.3	0.42	CACCGTATGAGCCCYTGTTGGAGGTGAAGAC
K01160	99	C	0.56	T	0.44	0.49	CATTGTGGCTGAACCAYGTCGCCCTCTTAATGGT
K01160	102	C	0.57	T	0.43	0.49	TGTGGCTGACCACGYGCCCTCTTAATGGTGTA
K01160	111	T	0.8	C	0.2	0.31	CCACGYGCCCTCTTA YGGTGTAAACTTGTAC
K01160	135	C	0.81	T	0.19	0.3	CTTGTACCAGTCTTAYGGTCCCCTCTGGCCAG
K01160	155	C	0.75	G	0.25	0.38	CCTCTGGCCAGTACASCCATGAATTTGATGG
K01160	279	C	0.42	G	0.58	0.49	GACAAACA TCGCTGTSTCTAAAACATAA CTTG
K01160	288	T	0.88	C	0.13	0.22	CGCTGTCTAAACAAYAACTTGAACAGTCTG
K01160	324	C	0.77	T	0.23	0.36	ACGCTCAA ACTCTACVYGTGCTACCAATGAG
K01160	378	A	0.71	G	0.29	0.41	CAAGTCTCCCGTGACRGTGGGTGAGCCCAAC
K01160	395	T	0.44	C	0.56	0.49	TGGGTCA GCCCAACA YCCTCATCTGCTTGT
K01160	444	A	0.6	C	0.4	0.48	TGTGGTCA A CATCACMTGGCTGAGCAATGGG
K01160	563	G	0.69	T	0.31	0.43	CTTCTGTGAGGAGAKTTATGACTGC AAAGGT
K01160	591	C	0.65	A	0.35	0.45	GGTGGAGCA CTGGGMCTGGACAAGCCTCTT
K01160	598	A	0.5	G	0.5	0.5	CAC TGGGGCCTGGACRAGCCTCTTCTGAAAC
K01160	636	C	0.88	A	0.13	0.22	GCCTGAGATTCCAGCMCCTATGTCAGAGCTC
K01160	669	C	0.81	T	0.19	0.3	AGAGACTGTGGTCTGYGCCCTGGGATTGTCT
K01160	678	A	0.46	G	0.54	0.5	GGTCTGGCCCTGGGRITGTCTGTGGGCCCTC
K01160	728	G	0.89	A	0.11	0.19	CTGTCTTCA TCATCCRAGGCTGCGTTCA GT
K01160	768	C	0.81	A	0.19	0.3	CAGACACCAAGGGCCMTTGTGAATCCCCATCC
K01160	789	T	0.6	G	0.4	0.48	AATCCCATCTGTGGA AKGGAAGGTGCATCGCC
K01160	835	C	0.93	A	0.07	0.13	AGAGTGGACTTGTCTAMATGACCTAGCA TTAT
K01160	945	C	0.75	T	0.25	0.38	TATCTGCTCAGAGCTYACAAATGCCCTTTGAA
K01160	993	T	0.31	C	0.69	0.43	TGATTTTTTCTTCTYAAGTGTACCTACTA
K01160	1013	T	0.07	A	0.93	0.13	GTTAACCTA TAAGAGWTGCTGGAGTAAGCC
K01160	1021	A	0.94	G	0.06	0.12	CTAAGAGTTGCCCTGGRGTAGCCACCCAGCT
K01160	1029	A	0.81	G	0.19	0.3	TGCCTGGAGTAAGCCRCGCCAGCTA CCTAAT
K01160	1095	T	0.94	A	0.06	0.12	AACAAATTCCTTTAWGAGATATATGTC AAA
K01160	1101	A	0.75	C	0.25	0.38	TTCCCTTTATGAGATMTATGTCAAAATTTTC
K01160	1153	G	0.88	C	0.13	0.22	TGAAACCGTGGCTAASAATTGGGAGACTCTC
K01160	1159	G	0.88	T	0.13	0.22	CGTGGCTAAGAAATTGKGAGACTCTCTTGTTT
K01160	1208	G	0.77	T	0.23	0.36	TTTACCAGATCAATTKTCAATGCCAGTAACA
L12168	1160	C	0.88	T	0.13	0.22	CATTACAGTAGATAAAYTGTAAGAAACTTGGC
L12168	1222	A	0.88	G	0.13	0.22	AGATAATCAACAGTARGGATGTCAAAGTTCA
L12168	1322	A	0.8	G	0.2	0.32	TTCCCTGGATTGTGARATAGTCAGTGGCCAAA

7

5
10
15
20
25
30
35
40
45
50
55

9

[illegible]

[illegible]

5
10
15
20
25
30
35
40
45
50
55

5
10
15
20
25
30
35
40
45
50
55

[illegible]

15

5

10

15

20

25

30

35

40

45

50

55

X56468	331	C	0.63	T	0.38	0.47	CA'CGAGCAGAAACACACACACCT'CCGACAAG
X56468	721	C	0.94	G	0.06	0.1	GGCTTTGATGAGGCSATTCGTGAACCTTGAT
X56468	755	T	0.89	G	0.11	0.2	CTGAATGAAGACTCAKAKACAAAGACGACCCC
X56468	781	G	0.71	A	0.29	0.41	CACCTCATCATGTCARITGCTTAGAGACAAC
X56468	865	C	0.58	A	0.42	0.49	GGCTGAATACTAAATMCATACAGGGGTGTCAT
X56468	879	A	0.73	C	0.27	0.39	TCCATACAGGGGTGTCMTCTCTTCTTCTTCA
X56468	1165	C	0.72	A	0.28	0.4	TTTAAAGTAATCTGAAMCAGTTCTGCAAGTGA
X56468	1676	A	0.83	G	0.17	0.28	AAATTAAAGTACACARGTCCAAGTCTAAAAAC
X57351	34	T	0.3	C	0.7	0.42	AGGACCTCCACACCAATACCAAGATGAGCCT
X57351	186	C	0.9	G	0.1	0.18	TGCATTGACAAATGSCAGGAAAGAGAAACT
X57351	318	G	0.88	A	0.13	0.22	TCTCTGTCAACAGCGGCCAGCCTCCCAACT
X57351	512	C	0.79	G	0.21	0.33	CTTCATGACCATTCGCTACTCGTGAACTCT
X57351	635	G	0.63	C	0.38	0.47	CTTCATGACCATTCGCTACTCGTGAACTCT
X57351	639	G	0.67	A	0.33	0.44	ATGACCATTCGCTACTCGTGAACTCT
X59892	135	C	0.94	T	0.06	0.1	AGTGACCCGCTCTCTCTCTCTCTCTCTCTCT
X59892	270	G	0.89	T	0.11	0.2	AGCTACAAAGCTGCKCGGGGAGGATTACA
X59892	1113	T	0.72	C	0.28	0.41	CCTAAACCAAGCCCTGCTGCTGCTGCTGCT
X59892	2051	G	0.94	T	0.06	0.1	AA'TGATAAGTAAAKCTTCTGTTATACATTT
X59892	2159	A	0.93	G	0.07	0.13	AGACTTGGGGTGAARCTTCCAGTTTACTG
X59932	175	A	0.56	T	0.44	0.49	TCCGTACAGAAATGTWTTGCCAAGTACAACT
X59932	204	C	0.94	T	0.06	0.12	CTTCCACGGCACTGCGAGCAGGACCTGCC
X59932	891	C	0.88	T	0.12	0.21	GGTCCAGCTCTCTGGGTGATCTGTGGAGGAG
X59932	973	A	0.91	C	0.09	0.16	GACTACCTGGGTCTMGGGTCTGGTCACTGC
X59932	1411	G	0.94	A	0.06	0.1	TGCTGGCACCTGGAACCCGCCATGGGCCCT
X59932	1563	G	0.5	A	0.5	0.5	TGCCCTGCTCACTGRGCCGAGCCTGAACT
X59932	1704	G	0.94	A	0.06	0.12	GAGGAAAGAGGCCACRAGCGGGAGGCAGCG
X59932	1788	T	0.67	G	0.33	0.44	ATTCCTTCTCTTTTGAGATTTTCTTCTCG
X59932	2114	T	0.41	C	0.59	0.48	CTTGTGAGATGGAATYGTAAATAAACACGCC
Z47055	146	T	0.88	C	0.12	0.21	TGGAGTACAATGCCAYTGGAGGCAAGTATAA
STSG-13126	33	T	0.88	C	0.13	0.22	TCCGGCCCTATCTCTCTCTCTCTCTCTCTCT
STSG-13353	52	T	0.86	C	0.14	0.24	GAAATAGGACCAATCYGAGGCTTTCTGTAGA
STSG-13369	63	G	0.93	C	0.07	0.13	GTGCTTGGTATATASAGGTAAATGTACT
STSG-13448	40	T	0.69	C	0.31	0.43	ATGACAGATTCAAGAYGATGCGCTCTCTGT
D00761	42	C	0.6	G	0.4	0.48	GCCATGTATTCGGCTCTGCGCAGAGACTTGG
D00761	514	C	0.94	G	0.06	0.1	TCAAAGGCTGGAGGCTSAGCAAGTGCCATGCT

GACAAATGGTGG
AACCGCAAGTT
ATTCATTTCTC
TGTGTTCTCOG

17

18

5
10
15
20
25
30
35
40
45
50
55

20

21

[illegible]

	X64364	897	G	A	0.37	0.47	GACGCTCCCTGTCCRCGTCTGGCGCCGCCGC
5	X64364	1390	G	A	0.08	0.15	ACCGAGGGCGACCCCRTCACAGCCTCAAGTC
10	X64364	1540	T	C	0.31	0.43	GGACGGCCCGGCTCTCYATAGCACCCAGGGCTC
15							
20							
25							
30							
35							
40							
45							
50							
55							

I. ANALYSIS of POLYMORPHISMS

A. PREPARATION of SAMPLES

- 5 [0023] Polymorphisms are detected in a target nucleic acid from the subject(s) being analyzed. Any suitable biological sample can be used for assay of genomic DNA. Convenient suitable tissue samples include whole blood, semen, saliva, tears, urine, fecal material, sweat, buccal, skin and hair. Pure red blood cells are not suitable. As those skilled in the art will appreciate, for assays of cDNA or mRNA, the tissue sample must be obtained from an organ in which the target nucleic acid is expressed, e.g., the liver for a target nucleic acid of a cytochrom P450.
- 10 [0024] Many known methods such as, for example those described below require amplification of the DNA from target samples. This can be accomplished by e.g., PCR. See, e.g., PCR Technology: Principles and Applications for DNA Amplification (ed. H.A. Erlich, Freeman Press, NY, NY, 1992); PCR Protocols: A Guide to Methods and Applications (eds. Innis, et al., Academic Press, San Diego, CA, 1990); Mattila et al., Nucleic Acids Res. 19, 4967 (1991); Eckert et al., PCR Methods and Applications 1, 17 (1991); PCR (eds. McPherson et al. IRL Press, Oxford); and U.S. Patent
- 15 4,683,202.
- [0025] Other suitable amplification methods include the ligase chain reaction (LCR) (e.g., Wu and Wallace, Genomics 4,560 (1989) and Landegren et al., Science 241, 1077 (1988)), transcription amplification (Kwoh et al., Proc. Natl. Acad. Sci. USA 86, 1173 (1989)), and self-sustained sequence replication (Guatelli et al., Proc. Nat. Acad. Sci. USA, 87, 1874 (1990)) and nucleic acid based sequence amplification (NABSA). The latter two amplification methods include
- 20 isothermal reactions based on isothermal transcription, which produce both single-stranded RNA (ssRNA) and double-stranded DNA (dsDNA) as the amplification products in a ratio of about 30 or 100 to 1, respectively.

B. DETECTION of NOVEL POLYMORPHISMS in TARGET DNA

- 25 [0026] There are two distinct types of analyses depending upon whether a given polymorphism has already been characterized. The first type of analysis is sometimes referred to as de novo characterization and compares target sequences in different subjects to identify points of variation, i.e., polymorphic sites. By analyzing a group of subjects representing the greatest ethnic diversity among human beings and the greatest breed and species variety of plants and animals, patterns characteristic of the most common alleles/haplotypes of a given locus can be identified, and the
- 30 frequencies of such patterns in the population can be determined. Additional allelic frequencies can be determined for subpopulations characterized by criteria such as, for example, geography, race, or gender. This de novo characterization of the polymorphisms of this invention is described in the EXAMPLES section of this description. The second type of analysis determines which form(s) of a characterized polymorphism are present in the subjects under test. Many suitable procedures exist and these are discussed immediately below.

1. ALLELE-SPECIFIC PROBES

- [0027] The design and use of allele-specific probes for analyzing polymorphisms is known in the art, e.g., Saiki et al., Nature 324, 163-166 (1986); Dattagupta, EP 235,726 and Saiki, WO 89/11548. Allele-specific probes can be
- 40 designed that hybridize to a segment of target DNA from one subject but not to the corresponding segment from another subject due to the presence of different polymorphic forms in their respective segments. Hybridization conditions should be suitably stringent so that there is a significant difference in hybridization intensity between alleles, preferably an essentially binary response, whereby the chosen probe hybridizes to only one of the alleles. Some probes are designed to hybridize to a segment of target DNA such that the polymorphic site aligns with a central position (e.g., in
- 45 a 15 mer at the 7 position; in a 16 mer, at either the 8 or 9 position) of the probe. This particular design of probe achieves useful discrimination in hybridization between different allelic forms.

- [0028] Allele-specific probes are often used in pairs, with one member of a pair showing a perfect match to a reference form of a target sequence and the other member showing a perfect match to a variant form. Several pairs of probes can then be immobilized on the same support for simultaneous analyses of multiple polymorphisms within the
- 50 same target sequence.

2. TILING ARRAYS

- [0029] Polymorphisms can also be identified by hybridization to nucleic acid arrays, (See, e.g., WO 95/11995). One
- 55 form of such arrays is described in the EXAMPLES section of this description in connection with de novo identification of polymorphisms. The same array or a different array can be used for analysis of characterized polymorphisms. The aforementioned WO 95/11995 also discloses subarrays that are optimized for the detection of variant forms of a pre-characterized polymorphism. Such subarrays contain probes designed to be complementary to a second reference

sequence, which is an allelic variant of the first reference sequence. The second group of probes is designed by the same principles as described in the EXAMPLES section of this description except that the probes exhibit complementarity to the second reference sequence. The inclusion of a second group (or further groups) can be particularly useful for analyzing short subsequences of the primary reference sequence in which multiple mutations are predicted to occur within a short distance commensurate with the length of the probes (i.e., two or more mutations within 9 to 21 bases).

3. ALLELE-SPECIFIC PRIMERS

[0030] An allele-specific primer hybridizes to a site on a target DNA overlapping a polymorphism and only primes amplification of an allelic form to which the primer exhibits perfect complementarity. (See, e.g., Gibbs, Nucleic Acid Res. 17, 2427-2448 (1989)). This primer is used in conjunction with a second primer which hybridizes at a distal site. Amplification proceeds from the two primers leading to a detectable product which signifies that the particular allelic form is present. A control is generally performed with a second pair of primers, one of which shows a single base mismatch at the polymorphic site and the other of which exhibits perfect complementarity the distal site. This single-base mismatch prevents amplification and, as such, no detectable product is formed. Most preferably, the mismatch is included in the 3'-most position of the oligonucleotide aligned with the polymorphism, i.e., the 3'-most position is the position most destabilizing to elongation from the primer. (See, e.g., WO 93/22456.)

4. DIRECT SEQUENCING

[0031] The direct analysis of the sequence of polymorphisms of this present invention can be accomplished by using either the dideoxy- chain termination method or the Maxam -Gilbert method (e.g., Sambrook et al., Molecular Cloning, A Laboratory Manual (2nd Ed., CSHP, New York 1989) and Zyskind et al., Recombinant DNA Laboratory Manual, (Acad. Press, 1988)).

5. DENATURING GRADIENT GEL ELECTROPHORESIS

[0032] Amplification products generated using PCR can be analyzed by the use of denaturing gradient gel electrophoresis. Different alleles can be identified based on the different sequence-dependent melting properties and electrophoretic migration of DNA in solution. (See, e.g., Erlich, ed., PCR Technology, Principles and Applications for DNA Amplification, (W.H. Freeman and Co, New York, 1992), Chapter 7.)

6. SINGLE STRAND CONFORMATION POLYMORPHISM

ANALYSIS

[0033] Alleles of target sequences can be differentiated using single-strand conformation polymorphism analysis, which identifies base differences by alteration in electrophoretic migration of single stranded PCR products. (See, e.g., Orita et al., Proc. Nat. Acad. Sci. 86, 2766-2770 (1989)). Amplified PCR products can be generated as described above, and heated or otherwise denatured, to form single stranded amplification products. Single-stranded nucleic acids may refold or form secondary structures which are partially dependent on the base sequence. The different electrophoretic mobilities of single-stranded amplification products can be related to base-sequence difference between alleles of the target sequences.

7. SINGLE BASE EXTENSION METHODS

[0034] Single base extension methods are described by e.g., US 5,846,710, US 6,004,744, US 5,888,819 and US 5,856,092. In brief, the methods work by hybridizing a primer that is complementary to a target strand such that the 3' end of the primer aligns with the base immediately 3' to the site of variation in the target strand but the primer does not include a base aligning with the site of variation. In other words, if the primer is formed from the complementary strand to the target strand, the primer includes sequence up to the base immediately 5' to the site of variation on the complementary strand. The hybridization is performed in the presence of one or more labelled nucleotides complementary to base(s) that may occupy the site of potential variation. For example, for a biallelic polymorphism two differentially labelled nucleotides can be used. In some methods, particularly, methods employing multiple differentially labelled nucleotides, the nucleotides are dideoynucleotides. Hybridization is performed under conditions permitting primer extension if a nucleotide complementary to a base occupying the site of variation in the target sequence is present. Extension incorporates a labelled nucleotide thereby generating a labelled extended primer. If multiple differentially labelled nucleotides are used and the target is heterozygous then multiple differentially labelled extended primers can be obtained.

Extended primers are detected providing an indication of which base(s) occupy the site of variation in the target polynucleotide.

III. METHODS of USE

[0035] The presence of polymorphic form(s) in a subject at one or more polymorphic sites is useful for, e.g., forensics, paternity testing, correlation of polymorphisms with phenotypic traits, and genetic mapping of phenotypic traits.

A. FORENSICS

[0036] Determination of which polymorphic forms occupy a set of polymorphic sites in an individual identifies a set of polymorphic forms that distinguishes the individual. (See, e.g., National Research Council, *The Evaluation of Forensic DNA Evidence* (Eds. Pollard et al., National Academy Press, DC, 1996)). Increasing the number of sites analyzed the probability that the set of polymorphic forms in one individual is the same as that in an unrelated individual. Preferably, where multiple sites are analyzed, the sites are unlinked. Thus, polymorphisms of this invention are often used in conjunction with polymorphisms in distal genes. Preferred polymorphisms for use in forensics are diallelic because the population frequencies of two polymorphic forms can usually be determined with greater accuracy than those of multiple polymorphic forms at multi-allelic loci.

[0037] As discussed above, the capacity to identify a distinguishing or unique set of forensic markers in an individual is useful for forensic analysis. For example, one can determine whether a blood sample from a suspect matches a blood or other tissue sample from a crime scene by determining whether the set of polymorphic forms occupying selected polymorphic sites is the same in the samples from suspect and the sample taken from the crime scene. Where the set of polymorphic markers taken from the crime scene does not the sample from the suspect, it can be concluded (barring experimental error) that the suspect is not the source of the sample taken from the crime scene. Where the set of markers does match, one can conclude that the DNA from the suspect is consistent with that found at the crime scene. Where frequencies of the polymorphic forms at the loci tested have been determined (e.g., by analysis of a suitable population of individuals), a statistical analysis can be performed to determine the probability that such a match of suspect and crime scene would occur merely by chance.

[0038] $p(ID)$ is the probability that two random individuals have the same polymorphic or allelic form at a given polymorphic site. In diallelic loci, four genotypes are possible: AA, AB, BA, and BB. Where alleles A and B occur in a haploid genome of the organism with frequencies x and y , the probability of each genotype in a diploid organism, as disclosed in WO 95/12607:

Homozygote: $p(AA) = x^2$.

Homozygote: $p(BB) = y^2 = (1-x)^2$.

Single Heterozygote: $p(AB) = p(BA) = xy = x(1-x)$.

Both Heterozygotes: $p(AB+BA) = 2xy = 2x(1-x)$.

[0039] The probability of identity at one locus (i.e., the probability that two individuals, picked at random from a population will have identical polymorphic forms at a given locus) is given by the equation:

$$p(ID) = (x^2)^2 + (2xy)^2 + (y^2)^2.$$

[0040] These calculations can be extended for any number of polymorphic forms at a given locus. For example, the probability of identity $p(ID)$ for a 3-allele system where the alleles have the frequencies in the population of x , y and z , respectively, is equal to the sum of the squares of the genotype frequencies:

$$p(ID) = x^4 + (2xy)^2 + (2yz)^2 + (2xz)^2 + z^4 + y^4.$$

[0041] In a locus of n alleles, the appropriate binomial expansion is used to calculate $p(ID)$ and $p(exc)$.

[0042] The cumulative probability of identity (cum $p(ID)$) for each of multiple unlinked loci is determined by multiplying the probabilities provided by each locus.:

$$\text{cum } p(ID) = p(ID1)p(ID2)p(ID3).... p(IDn).$$

[0043] The cumulative probability of non-identity for w loci (i.e., the probability that two random individuals will be different at 1 or more loci) is given by the equation:

$$\text{cum } p(\text{nonID}) = 1 - \text{cum } p(\text{ID}).$$

[0044] If several polymorphic loci are tested, the cumulative probability of non-identity for random individuals becomes very high (e.g., one in a billion). Such probabilities can be taken into account together with other evidence in determining the guilt or innocence of the suspect.

B. PATERNITY TESTING

[0045] The object of paternity testing is usually to determine whether a particular male is the father of a given child. In most cases, the mother of the child is known and thus, the mother's contribution to the child's genotype can be traced. Hence, paternity testing investigates whether the part of the child's genotype not attributable to the mother is consistent with that of the alleged father. Paternity testing can be performed by analyzing the sets of polymorphisms in both the alleged father and in his alleged child.

[0046] Where the set of polymorphisms in the child not attributable to the mother does not match the set of the alleged father, it can be concluded (barring experimental error) that the alleged father is not the child's father. Where the set of polymorphisms in the child not attributable to the mother does match the set of polymorphisms of the alleged father, a statistical calculation can be performed to determine the probability of a coincidental match.

[0047] The probability of parentage exclusion (representing the probability that a random male will have a polymorphic form at a given polymorphic site that makes him incompatible as the father) is given by the equation disclosed WO 95/12607:

$$p(\text{exc}) = xy(l-xy)$$

where x and y are the population frequencies of alleles A and B of a diallelic polymorphic site.

[0048] At a triallelic site $p(\text{exc}) = xy(l-xy) + yz(l-yz) + xz(l-xz) + 3xyz(l-xyz)$, where x, y and z are the respective population frequencies of alleles A, B, and C.

[0049] The probability of non-exclusion is:

$$p(\text{non-exc}) = l - p(\text{exc}).$$

[0050] The cumulative probability of non-exclusion (representing the value obtained when n loci are used) is thus:

$$\text{cum } p(\text{non-exc}) = p(\text{non-exc1})p(\text{non-exc2})p(\text{non-exc3})\dots p(\text{non-exc}n).$$

[0051] The cumulative probability of exclusion for z loci (representing the probability that a random male will be excluded)

$$\text{cum } p(\text{exc}) = 1 - \text{cum } p(\text{non-exc}).$$

[0052] Where several polymorphic loci are included in the analysis, the cumulative probability of exclusion of a random male is very high. This probability can be taken into account in assessing the liability of an alleged father whose polymorphic marker set matches the child's polymorphic marker set attributable to the child's father.

C. CORRELATION of POLYMORPHISMS with PHENOTYPIC TRAITS

[0053] The polymorphisms of this invention may contribute to the phenotype of an organism in different ways. As discussed above, some polymorphisms occur within a protein coding sequence and contribute to phenotype by affecting protein structure. The effect of such a change to the structure of a protein may be neutral, beneficial, or detrimental, or both beneficial and detrimental. For instance, a heterozygous sickle cell mutation confers resistance to malaria, but a homozygous sickle cell mutation is usually lethal. Other polymorphisms occur in noncoding regions and exert phenotypic effects indirectly via influence on, e.g., replication, transcription, and translation. Moreover, a single polymorphism may affect more than one phenotypic trait. Likewise, a single phenotypic trait may be affected by polymorphisms in different genes. Further, some polymorphisms predispose an individual to a distinct mutation that is causally related to a certain phenotype.

[0054] Phenotypic traits include diseases that have known (but hitherto unmapped) genetic components (e.g., agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome, muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial hypercholesterolemia, polycystic kidney disease, hereditary spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary hemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos syndrome,

osteogenesis imperfecta, and acute intermittent porphyria and the like). Phenotypic traits also include symptoms of, or susceptibility to, multifactorial diseases of which a component is, or may be, genetic, such as autoimmune diseases, inflammation, cancer, diseases of the nervous system, and infection by pathogenic microorganisms. Some examples of autoimmune diseases include rheumatoid arthritis, multiple sclerosis, diabetes (insulin-dependent and non-independent), systemic lupus erythematosus, and Graves disease. Some examples of cancers include cancers of the bladder, brain, breast, colon, esophagus, kidney, leukemia, liver, lung, oral cavity, ovary, pancreas, prostate, skin, stomach and uterus. Phenotypic traits also include characteristics such as longevity, appearance (e.g., baldness, obesity), strength, speed, endurance, fertility, and susceptibility or receptivity to particular drugs or therapeutic treatments.

[0055] Correlation of polymorphisms with phenotypic traits is performed for a population of subjects who have been tested for the presence or absence of a phenotypic trait of interest and for polymorphic markers sets. To perform such analysis, the presence or absence of a set of polymorphisms (i.e. a polymorphic set) is determined for a set of the subjects, some of whom exhibit a particular trait, and some of whom do not. The alleles of each polymorphism of the set are then reviewed to determine whether the presence or absence of a particular allele is associated with the phenotypic trait of interest. Correlation can be performed by using standard statistical methods such as a χ^2 -squared test and by noting any statistically significant correlations between the polymorphic form(s) and the phenotypic characteristics. For example, it might be found that the presence of allele A1 at polymorphism A correlates with heart disease. As a further example, it might be found that the combined presence of allele A1 at polymorphism A and allele B1 at polymorphism B correlates with increased milk production of a farm animal.

[0056] These correlations are useful in several ways. For example, where a strong correlation exists between a set of one or more polymorphic forms and a disease for which treatment is available, detection of the polymorphic form set in a subject, i.e., a human being or an animal may justify immediate administration of treatment, or at least the institution of regular monitoring of the subject. Detection of a polymorphic form correlated with a serious disease can assist in various types of decisions, for example, reproductive decision making. For instance, a female partner might elect to undergo in vitro fertilization to avoid the possibility of transmitting such a polymorphism from her male partner to her offspring.

[0057] Where a weaker yet statistically significant correlation exists between a polymorphic set and a human disease, immediate therapeutic intervention or monitoring may not be justified. Nevertheless, the patient can be motivated to begin simple life-style changes (e.g., diet, exercise) that can be accomplished at little cost to the patient but confer potential benefits in reducing the risk of conditions to which the patient may have increased susceptibility by virtue of variant alleles. Identification of a polymorphic set in a patient correlated with enhanced receptiveness to one of several treatment regimes for a disease indicates that this treatment regime should be followed.

[0058] For animals and plants, correlations between characteristics and phenotype are useful for e.g., breeding for desired characteristics. For example, Beitz et al., in US 5,292,639, disclose use of bovine mitochondrial polymorphisms in a breeding program to improve milk production in cows. To evaluate the effect of mitochondrial DNA (mtDNA) displacement loop (D-loop) sequence polymorphism on milk production, each cow was assigned a value of 1, where variant, or 0, where wildtype, with respect to a prototypical mtDNA sequence at each of the 17 locations considered. Each production trait was analyzed individually with the following animal model:

$$Y_{ijknp} = \mu + YS_i + P_j + X_k + \beta_1 + \dots \beta_{17} + PE_n + a_n + e_p$$

wherein: Y_{ijknp} is the milk, fat, fat percentage, solids-not-fat (SNF), SNF percentage, energy concentration, or lactation energy record; μ is an overall mean; YS_i is the effect common to all cows calving in year-season; X_k is the effect common to cows in either the high or average selection line; β_1 to β_{17} are the binomial regressions of production record on mtDNA D-loop sequence polymorphisms; PE_n is permanent environmental effect common to all records of cow n ; a_n is the effect of animal n and is composed of the additive genetic contribution of sire and dam breeding values and a Mendelian sampling effect; and e_p is a random residual. It was found that eleven of the seventeen polymorphisms tested influenced at least one production trait. Bovines having the best polymorphic forms for milk production at these eleven loci are selected to breed the next generation of the herd.

D. GENETIC MAPPING of PHENOTYPIC TRAITS

[0059] The previous section described the identification of correlations between phenotypic traits and polymorphisms that directly or indirectly contribute to those traits. The present section describes the identification of a physical linkage between a genetic locus associated with a trait of interest and polymorphic markers that are not associated with that trait, but rather are in physical proximity with the genetic locus responsible for the trait and co-segregate with it. Such analysis is useful for, e.g., mapping a genetic locus associated with a phenotypic trait to a chromosomal position, and thereby cloning the gene(s) responsible for the trait. (See, e.g., Lander et al., *Proc. Natl. Acad. Sci. (USA)* 83, 7353-7357 (1986); Lander et al., *Proc. Natl. Acad. Sci. (USA)* 84, 2363-2367 (1987); Donis-Keller et al., *Cell* 51, 319-337

(1987); Lander et al., *Genetics* 121, 185-199 (1989)). Genes localized by linkage can be cloned by a process known as directional cloning. (See, e.g., Wainwright, *Med. J. Australia* 159, 170-174 (1993); and Collins, *Nature Genetics* 1, 3-6 (1992)).

5 [0060] Linkage studies are generally performed on members of a family. Available members of the family are characterized for the presence or absence of a phenotypic trait and for a set of polymorphic markers. The distribution of polymorphic markers in an informative meiosis is then analyzed to determine which polymorphic markers co-segregate with a phenotypic trait. (See, e.g., Kerem et al., *Science* 245, 1073-1080 (1989); Monaco et al., *Nature* 316, 842 (1985); Yamoka et al., *Neurology* 40, 222-226 (1990); and Rossiter et al., *FASEB Journal* 5, 21-27 (1991)).

10 [0061] Linkage is analyzed by calculation of log of the odds (LOD) values. A LOD value is the relative likelihood of obtaining observed segregation data for a marker and a genetic locus when the two are located at a recombination fraction θ , versus the situation in which the two are not linked, and thus segregating independently (See, e.g., Thompson & Thompson, *Genetics in Medicine* (5th ed, W.B. Saunders Company, Philadelphia, 1991); and Strachan, "Mapping the human genome" in *The Human Genome* (BIOS Scientific Publishers Ltd, Oxford), Chapter 4). A series of likelihood ratios are calculated at various θ , ranging from $\theta = 0.0$ (coincident loci) to $\theta = 0.50$ (unlinked). Thus, the likelihood at a given value of θ is the probability of data where loci are linked at θ to the probability of data where loci are unlinked. The computed likelihoods are usually expressed as the \log_{10} of this ratio (i.e., a LOD value). For example, a LOD value of 3 indicates 1000:1 odds against an apparent observed linkage being a coincidence. The use of logarithms allows data collected from different families to be combined by simple addition. Computer programs are available for the calculation of LOD scores for differing values of θ (e.g., LIPED, MLINK (See e.g., Lathrop, *Proc. Nat. Acad. Sci. (USA)* 81, 3443-3446 (1984)). For any particular LOD score, a recombination fraction may be determined from mathematical tables. (See, e.g., Smith et al., *Mathematical tables for research workers in human genetics* (Churchill, London, 1961); and Smith, *Ann. Hum. Genet.* 32, 127-150 (1968)). The value of θ at which the LOD score is the highest is considered to be the best estimate of the recombination fraction. Positive LOD score values suggest that the two loci are linked, whereas negative LOD values suggest that linkage is less likely (at that value of θ) than the possibility that the two loci are unlinked. By convention, a combined LOD value of +3 or greater (equivalent to greater than 1000:1 odds in favor of linkage) is considered definitive evidence that the two loci are linked. Similarly, by convention, a negative LOD value of -2 or less is taken as definitive evidence against linkage of the two loci being compared. Negative linkage data are useful in excluding a chromosome or a segment thereof from consideration. The search then focuses on the remaining non-excluded chromosomal locations.

30

IV. MODIFIED POLYPEPTIDES and GENE SEQUENCES

[0062] This invention further provides variant forms of nucleic acids and corresponding proteins. The nucleic acids comprise one of the sequences described in TABLE 1, column 8, in which the polymorphic position is occupied by one of the alternative bases for that position. Some nucleic acid encode full-length variant forms of proteins. Similarly, variant proteins have the prototypical amino acid sequences encoded by the nucleic acid sequence shown in TABLE 1, column 8, (read so as to be in-frame with the full-length coding sequence of which it is a component) except at an amino acid encoded by a codon including one of the polymorphic positions shown in TABLE 1. That position is occupied by the amino acid coded by the corresponding codon in any of the alternative forms shown in TABLE 1.

40 [0063] Variant genes can be expressed in an expression vector in which a variant gene is operably linked to a native or other promoter. Usually, the promoter is a eukaryotic promoter for expression in a mammalian cell. The transcription regulation sequences typically include a heterologous promoter and optionally an enhancer that is recognized by the host. The selection of a promoter, for example trp, lac, phage, glycolytic enzyme and tRNA promoters, depends, in part, on the host selected. Commercially available expression vectors can be used in this invention. Vectors can include host-recognized replication systems, amplifiable genes, selectTABLE markers, host sequences useful for insertion into the host genome, and the like.

[0064] The means of introducing the expression construct into a host cell varies depending upon, for example, the particular construction and the target host. Suitable means include fusion, conjugation, transfection, transduction, electroporation or injection, as described, for example, in Sambrook, *supra*. A wide variety of prokaryotic and eukaryotic host cells can be employed for expression of the variant gene. Suitable host cells include, for example, bacteria such as *E. coli*, yeast, filamentous fungi, insect cells, mammalian cells, typically immortalized, e.g., mouse, CHO, human and monkey cell lines, as well as derivatives thereof. Preferred host cells are able to process the variant gene product to produce the mature polypeptide, where processing includes glycosylation, ubiquitination, disulfide bond formation, general post-translational modification, and the like.

50 [0065] The protein may be isolated by conventional means of protein biochemistry and purification to obtain a substantially pure product, i.e., 80, 95 or 99% free of cell component contaminants. (See, e.g., Jacoby, *Methods in Enzymology* Volume 104, Academic Press, New York (1984); Scopes, *Protein Purification, Principles and Practice*, 2nd Edition, Springer-Verlag, New York (1987); and Deutscher (ed), *Guide to Protein Purification, Methods in Enzymology*,

Vol. 182 (1990)). Where the protein is secreted, it can be isolated from the supernatant in which the host cell is grown. Where the protein is not secreted, the protein can be isolated from a lysate of the host cells.

[0066] This invention further provides transgenic nonhuman animals capable of expressing an exogenous variant gene and/or having one or both alleles of an endogenous variant gene inactivated. Expression of an exogenous variant gene is usually achieved by operably linking the gene to a promoter and optionally an enhancer, and microinjecting the construct into a zygote. (See, e.g., Hogan et al., "Manipulating the Mouse Embryo, A Laboratory Manual," Cold Spring Harbor Laboratory.) Inactivation of endogenous variant genes can be achieved by forming a transgene in which a cloned variant gene is inactivated by insertion of a positive selection marker. (See, e.g., Capecchi, *Science* 244, 1288-1292 (1989)). The transgene is then introduced into an embryonic stem cell, where it undergoes homologous recombination with an endogenous variant gene. Mice and other rodents are preferred animals. Such animals provide useful drug screening systems.

[0067] In addition to substantially full-length polypeptides expressed by variant genes, this present invention also includes biologically active fragments of the polypeptides, or analogs thereof, including organic molecules which simulate the interactions of the peptides. Biologically active fragments include any portion of the full-length polypeptide that confers a biological function on the variant gene product, including ligand and antibody binding. Ligand binding includes binding by nucleic acids, proteins or polypeptides, small biologically active molecules, or large cellular structures.

[0068] Polyclonal and/or monoclonal antibodies that specifically bind to variant gene products but not to corresponding prototypical gene products are also provided by this invention. Antibodies can be made by injecting mice or other animals with the variant gene product or synthetic peptide fragments thereof. Monoclonal antibodies are screened as are described, for example, in Harlow & Lane, *Antibodies, A Laboratory Manual*, Cold Spring Harbor Press, New York (1988); and Goding, *Monoclonal antibodies, Principles and Practice* (2d ed.) Academic Press, New York (1986). Monoclonal antibodies are tested for specific immunoreactivity with a variant gene product and lack of immunoreactivity to the corresponding prototypical gene product. These antibodies are useful in diagnostic assays for detection of the variant form, or as an active ingredient in a pharmaceutical composition.

V. KITS

[0069] This invention further provides kits comprising at least one allele-specific oligonucleotide as described above. Preferably, the kits contain one or more pairs of allele-specific oligonucleotides hybridizing to different forms of a polymorphism. In some kits, the allele-specific oligonucleotides are provided immobilized to a substrate. For example, the same substrate can comprise allele-specific oligonucleotide probes for detecting at least 10, 100 or all of the polymorphisms shown in TABLE 1. Optional additional components of the kit include, for example, restriction enzymes, reverse-transcriptase or polymerase, the substrate nucleoside triphosphates, means used to label (for example, an avidin-enzyme conjugate and enzyme substrate and chromogen where the label is biotin), and the appropriate buffers for reverse transcription, PCR, or hybridization reactions. Generally, the kit also contains instructions for carrying out the methods.

VI. Computer Systems For Storing Polymorphism Data

[0070] Fig. 1A depicts a block diagram of a computer system 10 suitable for implementing the present invention. Computer system 10 includes a bus 12 which interconnects major subsystems such as a central processor 14, a system memory 16 (typically RAM), an input/output (I/O) controller 18, an external device such as a display screen 24 via a display adapter 26, serial ports 28 and 30, a keyboard 32, a fixed disk drive 34 via a storage interface 35 and a floppy disk drive 36 operative to receive a floppy disk 38, and a CD-ROM (or DVD-ROM) device 40 operative to receive a CD-ROM 42. Many other devices can be connected such as a user pointing device, e.g., a mouse 44 connected via serial port 28 and a network interface 46 connected via serial port 30.

[0071] Many other devices or subsystems (not shown) may be connected in a similar manner. Also, it is not necessary for all of the devices shown in Fig. 1A to be present to practice the present invention, as discussed below. The devices and subsystems may be interconnected in different ways from that shown in Fig. 1A. The operation of a computer system such as that shown in Fig. 1A is well known. Databases storing polymorphism information according to the present invention can be stored, e.g., in system memory 16 or on storage media such as fixed disk 34, floppy disk 38, or CD-ROM 42. An application program to access such databases can be operably disposed in system memory 16 or sorted on storage media such as fixed disk 34, floppy disk 38, or CD-ROM 42.

[0072] Fig. 1B depicts the interconnection of computer system 10 to remote computers 48, 50, and 52. Fig. 1B depicts a network 54 interconnecting remote servers 48, 50, and 52. Network interface 46 provides the connection from client computer system 10 to network 54. Network 54 can be, e.g., the Internet. Protocols for exchanging data via the Internet and other networks are well known. Information identifying the polymorphisms described herein can be transmitted across network 54 embedded in signals capable of traversing the physical media employed by network 54.

[0073] Information identifying polymorphisms shown in Table 1 is represented in records, which optionally, are subdivided into fields. Each record stores information relating to a different polymorphisms in Table 1. Collectively, the records can store information relating to all of the polymorphisms in Table 1, or any subset thereof, such as 5, 10, 50, or 100 polymorphisms from Table 1. In some databases, the information identifies a base occupying a polymorphic position and the location of the polymorphic position. The base can be represented as a single letter code (i.e., A, C, G or T/U) present in a polymorphic form other than that in the reference allele. Alternatively, the base occupying a polymorphic site can be represented in IUPAC ambiguity code as shown in Table 1. The location of a polymorphic site can be identified as its position within one of the sequences shown in Table 1. For example, in the first sequence shown in Table 1, the polymorphic site occupies the A or C base. The position can also be identified by reference to, for example, a chromosome, and distance from known markers within the chromosome. In other databases, information identifying a polymorphism contains sequences of 10-100 bases shown in Table 1 or the complements thereof, including a polymorphic site. Preferably, such information records at least 10, 15, 20, or 30 contiguous bases of sequences including a polymorphic site.

EXAMPLES

[0074] The polymorphisms shown in TABLE 1 were identified by resequencing of target sequences from eight unrelated individuals of diverse ethnic and geographic backgrounds by hybridization to probes immobilized to microfabricated arrays. The strategy and principles for design and use of such arrays are generally described in WO 95/11995. The strategy provides arrays of probes for analysis of target sequences showing a high degree of sequence identity to the reference sequences of the fragments shown in TABLE 1, column 1. The reference sequences were sequence-tagged sites (STSs) developed in the course of the Human Genome Project (see, e.g., *Science* 270, 1945-1954 (1995); *Nature* 380, 152-154 (1996)). Most STS's ranged from 100 base pair (bp) to 300 bp in size.

[0075] A typical probe array used in this analysis has two groups of four sets of probes that respectively tile both strands of a reference sequence. A first probe set comprises a plurality of probes exhibiting perfect complementarity with one of the reference sequences. Each probe in the first probe set has an interrogation position that corresponds to a nucleotide in the reference sequence. That is, the interrogation position is aligned with the corresponding nucleotide in the reference sequence, where the probe and reference sequence are aligned to maximize complementarity between the two. For each probe in the first set, there are three corresponding probes from three additional probe sets. Thus, there are four probes corresponding to each nucleotide in the reference sequence. The probes from these three additional probe sets are identical to the corresponding probe from the first probe set except at the interrogation position, which occurs in the same position in each of the four corresponding probes from the four probe sets, and is occupied by a different nucleotide in the four probe sets. In the present analysis, probes were 25 nucleotides long. Arrays tiled for multiple different reference sequences were included on the same substrate.

[0076] Multiple target sequences from an individual were amplified from human genomic DNA using primers for the fragments. The amplified target sequences were fluorescently labeled during or after PCR. The labeled target sequences were hybridized with a substrate bearing immobilized arrays of probes. The amount of label bound to the probes was measured. Analysis of the pattern of label revealed the nature and position of differences between the target and reference sequence. For example, comparison of the intensities of four corresponding probes reveals the identity of a corresponding nucleotide in the target sequences aligned with the interrogation position of the probes. The corresponding nucleotide is the complement of the nucleotide occupying the interrogation position of the probe showing the highest intensity. (See, e.g., WO 95/11995). The existence of a polymorphism is also manifested by differences in normalized hybridization intensities of probes flanking the polymorphism where the probes hybridized to corresponding targets from different individuals. For example, relative loss of hybridization intensity in a "footprint" of probes flanking a polymorphism signals a difference between the target and the reference sequence (i.e., a polymorphism) (See, e.g., EP 717,113). Additionally, hybridization intensities for corresponding targets from different individuals can be classified into groups or clusters suggested by the data, not defined *a priori*, such that isolates in a give cluster tend to be similar and isolates in different clusters tend to be dissimilar. See US Patent Application No. 08/797,812, filed February 7, 1997. Hybridizations to samples from different individuals were performed separately. TABLE 1 summarizes the data obtained for target sequences in comparison with a reference sequence for the eight individuals tested.

[0077] From the foregoing, it is clear that the invention includes a number of general uses that can be expressed concisely as follows. This invention provides for the use of any of the nucleic acid segments described above in the diagnosis or monitoring of diseases, such as cancer, inflammation, heart disease, diseases of the CNS, and susceptibility to infection by microorganisms. This invention further provides for the use of any of the nucleic acid segments in the manufacture of a medicament for the treatment (includes, inter alia, preventative (e.g., prophylactic), palliative and curative treatment) of such diseases. This invention further provides for the use of any of the DNA segments as a pharmaceutical in any suitable dosage form.

[0078] The present invention is described in full, clear, concise and exact terms to enable those skilled in the art to

make and use this invention. In addition, EXAMPLES and illustrations are provided albeit solely for the purposes of even more clarity and understanding and, as such, do not limit this invention which is defined by the appendant claims. It will also be understood that certain changes and modifications that may be practiced are within the scope of the appendant claims.

5

SEQUENCE LISTING

10

<110> Affymetrix Inc.
 <120> Genetic Compositions and Methods
 <130> 52944
 <150> US 09/238,402
 <151> 1999-01-27
 <160> 632

15

20

<210> 1
 <211> 31
 <212> DNA
 <213> unknown
 <308> D42043
 <400> 1
 ACCAGCGAGG GGAGTRTATC CACCAAGCAG A

25

30

<210> 2
 <211> 31
 <212> DNA
 <213> unknown
 <308> D42043
 <400> 2
 GGGAAACTGT GTTTCRGGAC AGCAGCAGGA G

35

40

<210> 3
 <211> 31
 <212> DNA
 <213> unknown
 <308> D42043
 <400> 3
 GAGCATTCGG GTCAAWTCCA TGGACACCCT G

45

50

55

<210> 4
 <211> 31
 <212> DNA
 <213> unknown
 <308> D45248
 <400> 4
 GAGGGTGAAT GCCGTSAGA CCAAAGTGGA A

<210> 5
 <211> 31
 <212> DNA
 <213> unknown
 <308> M69066
 <400> 5
 ATGGTGCCTT CAAGASCTTC ACCAAACATA C


```

5      <210> 6
      <211> 31
      <212> DNA
      <213> unknown
      <308> M69066
      <400> 6
10     CTCCACAGCC ATCCCRGCCT TGGCATCTAG A

      <210> 7
      <211> 31
      <212> DNA
      <213> unknown
      <308> M69066
      <400> 7
15     GATTTTGGCC TCATTMCTTT ACCACCCCCA C

      <210> 8
      <211> 31
      <212> DNA
      <213> unknown
      <308> U56637
      <400> 8
20     AGCAAGTTTG TTCCARGTGA CCCATTGAGC T

      <210> 9
      <211> 31
      <212> DNA
      <213> unknown
      <308> X12671
      <400> 9
25     AGGGCGAAGG TAGGCWGGCA GATACGTTTC T

      <210> 10
      <211> 31
      <212> DNA
      <213> unknown
      <308> X12671
      <400> 10
30     CAGGTTCTAT TTGGAWTTTA TATACAACCT G

      <210> 11
      <211> 31
      <212> DNA
      <213> unknown
      <308> AB006782
      <400> 11
35     GAGATGGCCT TCAGCRGTTC CCAGGCTCCC T
50

```

5
 <210> 12
 <211> 31
 <212> DNA
 <213> unknown
 <308> AB006782
 <400> 12
 AGCAGGCCTG ATGGCWTCCC ACTGGCCTCC A

10
 <210> 13
 <211> 31
 <212> DNA
 <213> unknown
 <308> AB006782
 <400> 13
 CACTGGCCTC CACCAMCTGA CCAGAGTGTT C

20
 <210> 14
 <211> 31
 <212> DNA
 <213> unknown
 <308> AB006782
 <400> 14
 TGACCAGAGT GTTCTYTTCA GAGGACTGGC T

30
 <210> 15
 <211> 31
 <212> DNA
 <213> unknown
 <308> AB006782
 <400> 15
 GAGTGTTCTC TTCAGRGGAC TGGCTCCTTT C

40
 <210> 16
 <211> 31
 <212> DNA
 <213> unknown
 <308> HT3708
 <400> 16
 GGAAACCTCA TTGTTSTGTA CAAAGTACTA G

50
 <210> 17
 <211> 31
 <212> DNA
 <213> unknown
 <308> HT3708
 <400> 17
 TGAAACTCAT CTCATYACCC TTTTCTGGGT T

55

5
5
10
15
20
25
30
35
40
45
50
55

<210> 18
<211> 31
<212> DNA
<213> unknown
<308> J02621
<400> 18
AGAAACTAAA GAAGAYTTAC CTGCGGAAAA C

<210> 19
<211> 31
<212> DNA
<213> unknown
<308> J02621
<400> 19
AGAAGACTTA CCTGCRGAAA ACGGGGAAAC G

<210> 20
<211> 31
<212> DNA
<213> unknown
<308> J02621
<400> 20
GAGGAATATT TTTATSAACT ATTTTGTAAT G

<210> 21
<211> 31
<212> DNA
<213> unknown
<308> J02621
<400> 21
TTTTTATCAA CTATTYTGTA ATGCAAGTTT T

<210> 22
<211> 31
<212> DNA
<213> unknown
<308> J02621
<400> 22
TTGTTTATTT TTTGGYACAA CCAGAAAATA G

<210> 23
<211> 31
<212> DNA
<213> unknown
<308> J02621
<400> 23
TTTGGTACAA CCAGARAATA GTGTGGGATA T

5
 <210> 24
 <211> 31
 <212> DNA
 <213> unknown
 <308> J02621
 <400> 24
 TACAACCAGA AAATASTGTG GGATATTGAA T

10
 <210> 25
 <211> 31
 <212> DNA
 <213> unknown
 <308> J02621
 <400> 25
 GTGGGATATT GAATTRTGGG AGGCTCTGAC T

15
 <210> 26
 <211> 31
 <212> DNA
 <213> unknown
 <308> J02621
 <400> 26
 GAATTATGGG AGGCTYTGAC TGTCTCGGGT G

20
 <210> 27
 <211> 31
 <212> DNA
 <213> unknown
 <308> J02621
 <400> 27
 AATACAAAGC ATATTMAATG GCAATATGGA G

25
 <210> 28
 <211> 31
 <212> DNA
 <213> unknown
 <308> J02621
 <400> 28
 TAAATTACTT CTATTMCCAT GTTGTTTTTT A

30
 <210> 29
 <211> 31
 <212> DNA
 <213> unknown
 <308> J02621
 <400> 29
 TTACTTCTAT TACCAYGTTG TTTTtagta G

35
 <210> 29
 <211> 31
 <212> DNA
 <213> unknown
 <308> J02621
 <400> 29
 TTACTTCTAT TACCAYGTTG TTTTtagta G

40
 <210> 29
 <211> 31
 <212> DNA
 <213> unknown
 <308> J02621
 <400> 29
 TTACTTCTAT TACCAYGTTG TTTTtagta G

45
 <210> 29
 <211> 31
 <212> DNA
 <213> unknown
 <308> J02621
 <400> 29
 TTACTTCTAT TACCAYGTTG TTTTtagta G

50
 <210> 29
 <211> 31
 <212> DNA
 <213> unknown
 <308> J02621
 <400> 29
 TTACTTCTAT TACCAYGTTG TTTTtagta G

55

5
<210> 30
<211> 31
<212> DNA
<213> unknown
<308> J02621
<400> 30
TTACCATGTT GTTTTKTAGT AGAATTGTTT C

10
<210> 31
<211> 31
<212> DNA
<213> unknown
15
<308> J02621
<400> 31
GAATTGTGTG CACTCKGTAA CATCTTTGGT T

20
<210> 32
<211> 31
<212> DNA
<213> unknown
<308> J02621
<400> 32
25
TGTAGTGTAC GTGCTRTTGA GTTGTGAACT G

30
<210> 33
<211> 31
<212> DNA
<213> unknown
<308> K01160
<400> 33
GATGCTGGGG GCCCTYGCCC TGACCACCGT G

35
<210> 34
<211> 31
<212> DNA
<213> unknown
<308> K01160
40
<400> 34
CCCTGACCAC CGTGAYGAGC CCCTGTGGAG G

45
<210> 35
<211> 31
<212> DNA
<213> unknown
<308> K01160
50
<400> 35
CACCGTGATG AGCCCYTGTG GAGGTGAAGA C

55

EP 1 024 200 A2

5
 <210> 36
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 36
 CATTGTGGCT GACCAYGTCG CCTCTTATGG T

10
 <210> 37
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 37
 TGTGGCTGAC CACGTYGCCT CTTATGGTGT A

15
 <210> 38
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 38
 CCACGTCGCC TCTTAYGGTG TAAACTTGTA C

20
 <210> 39
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 39
 CTTGTACCAG TCTTAYGGTC CCTCTGGCCA G

25
 <210> 40
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 40
 CCTCTGGCCA GTACASCCAT GAATTTGATG G

30
 <210> 41
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 41
 GACAAACATC GCTGTSCTAA AACATAACTT G

35
 <210> 41
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 41
 GACAAACATC GCTGTSCTAA AACATAACTT G

40
 <210> 41
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 41
 GACAAACATC GCTGTSCTAA AACATAACTT G

45
 <210> 41
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 41
 GACAAACATC GCTGTSCTAA AACATAACTT G

50
 <210> 41
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 41
 GACAAACATC GCTGTSCTAA AACATAACTT G

55
 <210> 41
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 41
 GACAAACATC GCTGTSCTAA AACATAACTT G

5 <210> 42
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 42
 CGCTGTCCTA AAACAYAACT TGAACAGTCT G

10 <210> 43
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 43
 ACGCTCCAAC TCTACYGCTG CTACCAATGA G

15 <210> 44
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 44
 CAAGTCTCCC GTGACRCTGG GTCAGCCCAA C

20 <210> 45
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 45
 TGGGTCAGCC CAACAYCCTC ATCTGTCTTG T

25 <210> 46
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 46
 TGTGGTCAAC ATCACMTGGC TGAGCAATGG G

30 <210> 47
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 47
 CTTCTGCTGA GGAGAKTTAT GACTGCAAGG T

35 <210> 48
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 48
 CTTCTGCTGA GGAGAKTTAT GACTGCAAGG T

40 <210> 49
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 49
 CTTCTGCTGA GGAGAKTTAT GACTGCAAGG T

45 <210> 50
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 50
 CTTCTGCTGA GGAGAKTTAT GACTGCAAGG T

50 <210> 51
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 51
 CTTCTGCTGA GGAGAKTTAT GACTGCAAGG T

55 <210> 52
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 52
 CTTCTGCTGA GGAGAKTTAT GACTGCAAGG T

5
10
15
20
25
30
35
40
45
50
55

<210> 48
<211> 31
<212> DNA
<213> unknown
<308> K01160
<400> 48
GGTGGAGCAC TGGGGMCTGG ACAAGCCTCT T

<210> 49
<211> 31
<212> DNA
<213> unknown
<308> K01160
<400> 49
CACTGGGGCC TGGACRAGCC TCTTCTGAAA C

<210> 50
<211> 31
<212> DNA
<213> unknown
<308> K01160
<400> 50
GCCTGAGATT CCAGCMCCTA TGTCAGAGCT C

<210> 51
<211> 31
<212> DNA
<213> unknown
<308> K01160
<400> 51
AGAGACTGTG GTCTGYGCCC TGGGATTGTC T

<210> 52
<211> 31
<212> DNA
<213> unknown
<308> K01160
<400> 52
GGTCTGCGCC CTGGGRTTGT CTGTGGGCCT C

<210> 53
<211> 31
<212> DNA
<213> unknown
<308> K01160
<400> 53
CTGTCTTCAT CATCCRAGGC CTGCGTTCAG T

5
5
10
15
20
25
30
35
40
45
50
55

<210> 54
<211> 31
<212> DNA
<213> unknown
<308> K01160
<400> 54
CAGACACCAA GGGCCMTTGT GAATCCCATC C

<210> 55
<211> 31
<212> DNA
<213> unknown
<308> K01160
<400> 55
AATCCCATCC TGGAAGGAA GGTGCATCGC C

<210> 56
<211> 31
<212> DNA
<213> unknown
<308> K01160
<400> 56
AGAGTGGACT TGCTAMATGA CCTAGCATTA T

<210> 57
<211> 31
<212> DNA
<213> unknown
<308> K01160
<400> 57
TATCTGCTCA GAGCTYACAA ATGCCTTTGA A

<210> 58
<211> 31
<212> DNA
<213> unknown
<308> K01160
<400> 58
TGATTTTTTT CTTCTYAAGT GTTACCTACT A

<210> 59
<211> 31
<212> DNA
<213> unknown
<308> K01160
<400> 59
GTTACCTACT AAGAGWTGCC TGGAGTAAGC C

EP 1 024 200 A2

5
 <210> 60
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 60
 CTAAGAGTTG CCTGGRGTAA GCCACCCAGC T

10
 <210> 61
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 61
 TGCCTGGAGT AAGCCRCCCA GCTACCTAAT T

15
 <210> 62
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 62
 AACAAATTCC CTTTAWGAGA TATATGTCAA A

20
 <210> 63
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 63
 TTCCCTTTAT GAGATMTATG TCAAATTTT C

25
 <210> 64
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 64
 TGAAACCGTG GCTAASAATT GGGAGACTCT C

30
 <210> 65
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 65
 CGTGGCTAAG AATTGKGAGA CTCTCTTGTT T

35
 <210> 66
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 66
 TTTTCTTCTT TTTTCTTCTT T

40
 <210> 67
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 67
 TTTTCTTCTT TTTTCTTCTT T

45
 <210> 68
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 68
 TTTTCTTCTT TTTTCTTCTT T

50
 <210> 69
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 69
 TTTTCTTCTT TTTTCTTCTT T

55
 <210> 70
 <211> 31
 <212> DNA
 <213> unknown
 <308> K01160
 <400> 70
 TTTTCTTCTT TTTTCTTCTT T

EP 1 024 200 A2

```

5      <210> 66
      <211> 31
      <212> DNA
      <213> unknown
      <308> K01160
      <400> 66
      TTTACCAGAT CATTTKTCAT GTCCAGTAAC A
10
      <210> 67
      <211> 31
      <212> DNA
      <213> unknown
      <308> L12168
      <400> 67
      CATTACAGTA GATAAYTGTA AGAAACTTGG C
15
      <210> 68
      <211> 31
      <212> DNA
      <213> unknown
      <308> L12168
      <400> 68
      AGATAATCAA CAGTARGGAT GTCAAAGTTC A
20
      <210> 69
      <211> 31
      <212> DNA
      <213> unknown
      <308> L12168
      <400> 69
      TTCCCTGGAT TGTGARATAG TCAGTGCCAA A
25
      <210> 70
      <211> 31
      <212> DNA
      <213> unknown
      <308> L12168
      <400> 70
      GCTTCTCTGC TCTGARAAGC ACAGCTACCT G
30
      <210> 71
      <211> 31
      <212> DNA
      <213> unknown
      <308> L12168
      <400> 71
      GAGCATGAAG GTAGCRGAAG CTGGTGTGTT T
35

```

5
<210> 72
<211> 31
<212> DNA
<213> unknown
<308> M15661
<400> 72
TGTATGCCCCA GGGAARGAGG CGCTATGATC G

10
<210> 73
<211> 31
<212> DNA
<213> unknown
15 <308> M29064
<400> 73
TAATGCAGAA GTAAGRAAGG CTTTGTCTAG A

20
<210> 74
<211> 31
<212> DNA
<213> unknown
<308> M37766
<400> 74
25 GCTCTGGAAT TGCTAYTGCT GCCTCTGTCA C

30
<210> 75
<211> 31
<212> DNA
<213> unknown
<308> M37766
<400> 75
AGCATTCAAG GTCACYTGGT ACATATGACC G

35
<210> 76
<211> 31
<212> DNA
<213> unknown
<308> M37766
40 <400> 76
TCTAAGGTCC AGAAASAGGA CAACAGCACC T

45
<210> 77
<211> 31
<212> DNA
<213> unknown
<308> M37766
50 <400> 77
CCCTTATGCC ACATAWTTAC TCCAGGTGTT A

55

5
<210> 78
<211> 31
<212> DNA
<213> unknown
<308> M37766
<400> 78
AAACTTCAAG GCCAGMAGAT CTTGCCTGTT G

10
<210> 79
<211> 31
<212> DNA
<213> unknown
15 <308> U05340
<400> 79
TGCGCCAGAG GGTTAYCAGA ACAGACTGAA A

20
<210> 80
<211> 31
<212> DNA
<213> unknown
<308> U05340
25 <400> 80
CCAGAGGGTT ATCAGRACAG ACTGAAAGTA C

30
<210> 81
<211> 31
<212> DNA
<213> unknown
<308> U05340
<400> 81
CTCCTGGCTC CAGCCSGAAG ACCTGCCGTT A

35
<210> 82
<211> 31
<212> DNA
<213> unknown
40 <308> U05340
<400> 82
TGTCCAGTGG TTCACRTTCT GGCCACATCC A

45
<210> 83
<211> 31
<212> DNA
<213> unknown
<308> U05340
50 <400> 83
GGTAGCAGAA CACCAYGTGG CCACACTGAG T

55

5
5
10
15
20
25
30
35
40
45
50
55

<210> 84
<211> 31
<212> DNA
<213> unknown
<308> U05340
<400> 84
GTGGGCTGCG CTGGGYCCCA GATGGACGAC A

<210> 85
<211> 31
<212> DNA
<213> unknown
<308> U05340
<400> 85
TTTGGCCAGT GGTGGYAATG ATAAGTTGGT C

<210> 86
<211> 31
<212> DNA
<213> unknown
<308> U05340
<400> 86
TGATAACTTG GTCAAWGTGT GGCCTAGTGC T

<210> 87
<211> 31
<212> DNA
<213> unknown
<308> U05340
<400> 87
GGCAGTCCAA TGTCCYGGCA ACAGGAGGGG G

<210> 88
<211> 31
<212> DNA
<213> unknown
<308> U05340
<400> 88
AGTGATCGAC ACATTYGCAT CTGGAATGTG T

<210> 89
<211> 31
<212> DNA
<213> unknown
<308> U05340
<400> 89
CATTCGCATC TGGAAYGTGT GCTCTGGGGC C

5
 <210> 90
 <211> 31
 <212> DNA
 <213> unknown
 <308> U05340
 <400> 90
 GTGCCGTGGA TGCCWTTCC CAGGTGTGCT C

10
 <210> 91
 <211> 31
 <212> DNA
 <213> unknown
 <308> U05340
 <400> 91
 CCTCTGGTCT CCCAYTACA AGGAGCTCAT C

20
 <210> 92
 <211> 31
 <212> DNA
 <213> unknown
 <308> U05340
 <400> 92
 CATCCCGGGT CCTGASTCTG ACCATGAGCC C

30
 <210> 93
 <211> 31
 <212> DNA
 <213> unknown
 <308> U05340
 <400> 93
 CAGCAGATGA GACCCYGAGG CTATGGCGCT G

35
 <210> 94
 <211> 31
 <212> DNA
 <213> unknown
 <308> U05340
 <400> 94
 GCAGCCAAAA GCAGCYTCAT CCACCAAGGC A

40
 <210> 95
 <211> 31
 <212> DNA
 <213> unknown
 <308> U43522
 <400> 95
 GACTGCAATG TGCCGRTCTT AGCTGCTGCC T

55

EP 1 024 200 A2

5
 <210> 96
 <211> 31
 <212> DNA
 <213> unknown
 <308> U43522
 <400> 96
 GTCCGAGCCC CTGAGYCGAG TAAAGTTGGG C

10
 <210> 97
 <211> 31
 <212> DNA
 <213> unknown
 <308> U43522
 <400> 97
 AGTAAAGTTG GGCACRTTAC GCCGGCCTGA A

15
 <210> 98
 <211> 31
 <212> DNA
 <213> unknown
 <308> U43522
 <400> 98
 CTTCAATCCT GGGAARAAC TCAAAC TGGT C

20
 <210> 99
 <211> 31
 <212> DNA
 <213> unknown
 <308> U43522
 <400> 99
 ACTGGCTGCA CCCACMGATG ACAGTGGGTG A

25
 <210> 100
 <211> 31
 <212> DNA
 <213> unknown
 <308> U43522
 <400> 100
 GCACCCACAG ATGACRGTGG GTGAGGTGCA G

30
 <210> 101
 <211> 31
 <212> DNA
 <213> unknown
 <308> U43522
 <400> 101
 AACCCTGCGA AGGCCYGGAG GTCCACAGTA T

35
 <210> 102
 <211> 31
 <212> DNA
 <213> unknown
 <308> U43522
 <400> 102
 AACCCTGCGA AGGCCYGGAG GTCCACAGTA T

40
 <210> 103
 <211> 31
 <212> DNA
 <213> unknown
 <308> U43522
 <400> 103
 AACCCTGCGA AGGCCYGGAG GTCCACAGTA T

45
 <210> 104
 <211> 31
 <212> DNA
 <213> unknown
 <308> U43522
 <400> 104
 AACCCTGCGA AGGCCYGGAG GTCCACAGTA T

50
 <210> 105
 <211> 31
 <212> DNA
 <213> unknown
 <308> U43522
 <400> 105
 AACCCTGCGA AGGCCYGGAG GTCCACAGTA T

55
 <210> 106
 <211> 31
 <212> DNA
 <213> unknown
 <308> U43522
 <400> 106
 AACCCTGCGA AGGCCYGGAG GTCCACAGTA T

5
<210> 102
<211> 31
<212> DNA
<213> unknown
<308> U43522
<400> 102
TGTCTACACA AATCAYAAAG GGGAGAAAAT C

10
<210> 103
<211> 31
<212> DNA
<213> unknown
15 <308> U43522
<400> 103
TTTATATGAA TGATAMGTCC CCATTGACGC C

20
<210> 104
<211> 31
<212> DNA
<213> unknown
<308> U43522
25 <400> 104
TAACCTGGAC CGGACYGATG ACCTGGTGTA C

30
<210> 105
<211> 31
<212> DNA
<213> unknown
<308> U43522
35 <400> 105
CAAGATGCGG CTGGCRCAGC AGAACGCCGT G

40
<210> 106
<211> 31
<212> DNA
<213> unknown
<308> U43522
45 <400> 106
CGCCGTGACC TCCCTRAGTG AGGAGTGCAA G

50
<210> 107
<211> 31
<212> DNA
<213> unknown
<308> U43522
55 <400> 107
GATGCTGACG GCTTCRCACA CCCTGGCTGT G

EP 1 024 200 A2

5
 <210> 108
 <211> 31
 <212> DNA
 <213> unknown
 <308> U43522
 <400> 108
 GTGGGTCTTC CAGGGRGAAG GCCAAGGGGA G

10
 <210> 109
 <211> 31
 <212> DNA
 <213> unknown
 <308> U43522
 <400> 109
 CCCCTAGCTT TATATRTGGA CATGGCAGGC C

15
 <210> 110
 <211> 31
 <212> DNA
 <213> unknown
 <308> U46692
 <400> 110
 CAGAGCGTGC ACTTGYGATC CTAAAATAAG C

20
 <210> 111
 <211> 31
 <212> DNA
 <213> unknown
 <308> U66711
 <400> 111
 CAAGTCCCAT GGACAKGCTG ACAGGGTCCC C

25
 <210> 112
 <211> 31
 <212> DNA
 <213> unknown
 <308> U66711
 <400> 112
 CTGTGTACGT GGGTGKGCAG TGCACGTGAG A

30
 <210> 113
 <211> 31
 <212> DNA
 <213> unknown
 <308> X05276
 <400> 113
 AACTTGCCCA GGCCARAGAA GAGAACGTGG G

35
 <210> 114
 <211> 31
 <212> DNA
 <213> unknown
 <308> U66711
 <400> 114
 CTGTGTACGT GGGTGKGCAG TGCACGTGAG A

40
 <210> 115
 <211> 31
 <212> DNA
 <213> unknown
 <308> U66711
 <400> 115
 CTGTGTACGT GGGTGKGCAG TGCACGTGAG A

45
 <210> 116
 <211> 31
 <212> DNA
 <213> unknown
 <308> U66711
 <400> 116
 CTGTGTACGT GGGTGKGCAG TGCACGTGAG A

50
 <210> 117
 <211> 31
 <212> DNA
 <213> unknown
 <308> U66711
 <400> 117
 CTGTGTACGT GGGTGKGCAG TGCACGTGAG A

55
 <210> 118
 <211> 31
 <212> DNA
 <213> unknown
 <308> U66711
 <400> 118
 CTGTGTACGT GGGTGKGCAG TGCACGTGAG A

5
<210> 114
<211> 31
<212> DNA
<213> unknown
<308> X05276
<400> 114
TTCGCTTTGC TGGAARTGTC AAGCAAATTA T

10
<210> 115
<211> 31
<212> DNA
<213> unknown
<308> X05276
<400> 115
CCCCACACTT TGAAGMATAC AGACCCCAGT A

15
<210> 116
<211> 31
<212> DNA
<213> unknown
<308> X05276
<400> 116
TGCAGGACAA ACATCYTCTC AAGCAGTCAA C

20
<210> 117
<211> 31
<212> DNA
<213> unknown
<308> X05276
<400> 117
TCTCAAGCAG TCAACSTAGA ATGCTTGGGA A

25
<210> 118
<211> 31
<212> DNA
<213> unknown
<308> X05276
<400> 118
TTCGAGACCA GCCCARCCAA CATGGCGAAA C

30
<210> 119
<211> 31
<212> DNA
<213> unknown
<308> X60221
<400> 119
AGAATGCAG CCTTCYTAGG TCCAGGGGTA T

35
<210> 119
<211> 31
<212> DNA
<213> unknown
<308> X60221
<400> 119
AGAATGCAG CCTTCYTAGG TCCAGGGGTA T

40
<210> 119
<211> 31
<212> DNA
<213> unknown
<308> X60221
<400> 119
AGAATGCAG CCTTCYTAGG TCCAGGGGTA T

45
<210> 119
<211> 31
<212> DNA
<213> unknown
<308> X60221
<400> 119
AGAATGCAG CCTTCYTAGG TCCAGGGGTA T

50
<210> 119
<211> 31
<212> DNA
<213> unknown
<308> X60221
<400> 119
AGAATGCAG CCTTCYTAGG TCCAGGGGTA T

55

5
10
15
20
25
30
35
40
45
50
55

<210> 120
<211> 31
<212> DNA
<213> unknown
<308> X60221
<400> 120
TTTCTTTATC CTAAARCTGG TGTAACAGGA C

<210> 121
<211> 31
<212> DNA
<213> unknown
<308> X60221
<400> 121
AACAGGACCC TATGTRCTCG GAACTGGGCT T

<210> 122
<211> 31
<212> DNA
<213> unknown
<308> X60221
<400> 122
CCTTTTTGAT GTGCARAGGA ATAACATTGC T

<210> 123
<211> 31
<212> DNA
<213> unknown
<308> X60221
<400> 123
GGACTATCAT ATATCYGTGC AGAACATGAT G

<210> 124
<211> 31
<212> DNA
<213> unknown
<308> Y00097
<400> 124
ACTCTGCTGA AGCTGKCTGG GGGAGATGAT G

<210> 125
<211> 31
<212> DNA
<213> unknown
<308> Y00097
<400> 125
AATAGCAGAC ACACCYAGTG GAGACAAAAC T

5
 <210> 126
 <211> 31
 <212> DNA
 <213> unknown
 <308> Y00097
 <400> 126
 ATTCATTGAG AAATAYGACA AGTCTCTCCA C

10
 <210> 127
 <211> 31
 <212> DNA
 <213> unknown
 <308> Y00097
 <400> 127
 AATATGACAA GTCTCYCCAC CAAGCCATTG A

15
 <210> 128
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z48950
 <400> 128
 AAGCGCTCCC TCTACYGGCG GGGTGAAGAA G

20
 <210> 129
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z48950
 <400> 129
 CTCTAGCTGA TAATGSAAAC ACTAACTGGG G

25
 <210> 130
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z48950
 <400> 130
 CTCTGCTGAG GAAACRGTAC CGAAGTTCTT T

30
 <210> 131
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z49107
 <400> 131
 CTGGGAGGGC TGTACYCATC CAAGTCCATC C

35
 <210> 132
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z49107
 <400> 132
 CTGGGAGGGC TGTACYCATC CAAGTCCATC C

40
 <210> 133
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z49107
 <400> 133
 CTGGGAGGGC TGTACYCATC CAAGTCCATC C

45
 <210> 134
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z49107
 <400> 134
 CTGGGAGGGC TGTACYCATC CAAGTCCATC C

50
 <210> 135
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z49107
 <400> 135
 CTGGGAGGGC TGTACYCATC CAAGTCCATC C

55
 <210> 136
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z49107
 <400> 136
 CTGGGAGGGC TGTACYCATC CAAGTCCATC C

EP 1 024 200 A2

5
 <210> 132
 <211> 31
 <212> DNA
 <213> unknown
 <308> D11086
 <400> 132
 CCTCACTCTG CATTAWTGGT ACAAGAACTC G

10

15
 <210> 133
 <211> 31
 <212> DNA
 <213> unknown
 <308> D11086
 <400> 133
 GCTAGAACTG AACTGSAACA ACAGATTCTT G

20

25
 <210> 134
 <211> 31
 <212> DNA
 <213> unknown
 <308> D11086
 <400> 134
 GATTGATTAT CAGCCWTCTC TGTGTGTATT T

30

35
 <210> 135
 <211> 31
 <212> DNA
 <213> unknown
 <308> D11086
 <400> 135
 TGTGCCCCCA TGTAASCACC CCTTCATTTG G

40

45
 <210> 136
 <211> 31
 <212> DNA
 <213> unknown
 <308> D38076
 <400> 136
 CCGCCAAGGA CACTCRTGAG GACCATGATA C

50

55
 <210> 137
 <211> 31
 <212> DNA
 <213> unknown
 <308> D38076
 <400> 137
 AGAATACAGA CGAGTMCAAC CATGACCCTC A

5
<210> 138
<211> 31
<212> DNA
<213> unknown
<308> D38076
<400> 138
AACTTTTAA AATGCRGGCA AACTGTTCC G

10
<210> 139
<211> 31
<212> DNA
<213> unknown
15 <308> D38076
<400> 139
CGGGCAAAC TGTTCYGATT TGCCTCTGAG A

20
<210> 140
<211> 31
<212> DNA
<213> unknown
25 <308> D38076
<400> 140
GTTCCGATTT GCCTCWGAGA ACGATCTCCC A

30
<210> 141
<211> 31
<212> DNA
<213> unknown
35 <308> D38076
<400> 141
TCCGCCTCCT CATGCRGAGG GACAAGACCC T

40
<210> 142
<211> 31
<212> DNA
<213> unknown
45 <308> D38076
<400> 142
GGACAAGACC CTGAATCT GTGCCAACCA C

50
<210> 143
<211> 31
<212> DNA
<213> unknown
55 <308> D38076
<400> 143
GGAGCTGAAG CCCAAMGCAG GTAGCGACCG T

5
5
10
15
20
25
30
35
40
45
50
55

<210> 144
<211> 31
<212> DNA
<213> unknown
<308> D38076
<400> 144
GCCCAACGCA GGTAGYGACC GTGCCTGGGT C

<210> 145
<211> 31
<212> DNA
<213> unknown
<308> D38076
<400> 145
GGTCTGGAAC ACCCAYGCTG ACTTCGCCGA C

<210> 146
<211> 31
<212> DNA
<213> unknown
<308> D38076
<400> 146
GCTGACTTCG CCGACSAGTG CCCCAGCCA G

<210> 147
<211> 31
<212> DNA
<213> unknown
<308> D38076
<400> 147
GAGCTGCTGG CCATCYGCTT CCTGAATGCT G

<210> 148
<211> 31
<212> DNA
<213> unknown
<308> D38076
<400> 148
GTTTGAAGAA TGCAGRAAAG AGATCGAAGA G

<210> 149
<211> 31
<212> DNA
<213> unknown
<308> D38076
<400> 149
AAAAATGATC ATGCCRAAAA AGTGGCGGAA A

5
<210> 150
<211> 31
<212> DNA
<213> unknown
<308> D38076
<400> 150
CCTCTTTTTC GGTTTRTTTT TATTCTTTCA T

10
<210> 151
<211> 31
<212> DNA
<213> unknown
<308> D87017
15
<400> 151
GAAGCTACAG CTGCCRGGTC ACGCATGAAG G

20
<210> 152
<211> 31
<212> DNA
<213> unknown
<308> L06132
25
<400> 152
TACAGAGAAA TGGAAYACCG ACAATACACT A

30
<210> 153
<211> 31
<212> DNA
<213> unknown
<308> L06132
35
<400> 153
CGAGGGCTGG CTGGCYGGCT ACCAGATGAA T

40
<210> 154
<211> 31
<212> DNA
<213> unknown
<308> L06132
45
<400> 154
CTATTTTGCA GCATASCTAC CTCAGAATT T

50
<210> 155
<211> 31
<212> DNA
<213> unknown
<308> L06132
55
<400> 155
GGTACAGAAG AAACCYATTT CCAAAAAAGG T

5
5
10
15
20
25
30
35
40
45
50
55

<210> 156
<211> 31
<212> DNA
<213> unknown
<308> L06132
<400> 156
TCTAGAAATG GCTGCRAGTG GAAGCGGATA A

<210> 157
<211> 31
<212> DNA
<213> unknown
<308> L06132
<400> 157
AATGGCTGCA AGTGGWAGCG GATAATATGT A

<210> 158
<211> 31
<212> DNA
<213> unknown
<308> L06132
<400> 158
AATCGAACCT TGGTTYCCTA ACCCTAATTG A

<210> 159
<211> 31
<212> DNA
<213> unknown
<308> L06132
<400> 159
ACCTTGGTTC CCTAAYCCTA ATTGATGAGA G

<210> 160
<211> 31
<212> DNA
<213> unknown
<308> L06132
<400> 160
TGGTTCCTA ACCCTRATTG ATGAGAGGCT C

<210> 161
<211> 31
<212> DNA
<213> unknown
<308> L06132
<400> 161
GGTTCCTAA CCCTAMTTGA TGAGAGGCTC G

5
<210> 162
<211> 31
<212> DNA
<213> unknown
<308> L06132
<400> 162
GTTCCCTAAC CCTAAYTGAT GAGAGGCTCG C

10
<210> 163
<211> 31
<212> DNA
<213> unknown
15
<308> L06132
<400> 163
TTTTTAGACA GATCTYCATG ACCTGTTCCC A

20
<210> 164
<211> 31
<212> DNA
<213> unknown
<308> L06132
25
<400> 164
CAGCTATAAG CTTGGRAGTG TCTGTGTGAT T

30
<210> 165
<211> 31
<212> DNA
<213> unknown
<308> L06132
<400> 165
AATCACATGG TGACARCACT CAGAATCTAA A

35
<210> 166
<211> 31
<212> DNA
<213> unknown
40
<308> L06132
<400> 166
ACTCAATTTG TTTTTYAGCA GTTTAATGGG T

45
<210> 167
<211> 31
<212> DNA
<213> unknown
<308> M19311
50
<400> 167
ACTTCGCCAT GTGATKACAA ACCTTGAGAG G

55

5
<210> 168
<211> 31
<212> DNA
<213> unknown
<308> M19311
<400> 168
GCATGTGGCT TACTCYGGCT ATATCTAAGC C

10
<210> 169
<211> 31
<212> DNA
<213> unknown
15
<308> M19311
<400> 169
CACATCTAAA CTTAGRTGGA GTTGGTCAAA T

20
<210> 170
<211> 31
<212> DNA
<213> unknown
<308> M22760
25
<400> 170
TTCGAATCCT AGAGGYTGTT AAGGACAAAG C

30
<210> 171
<211> 31
<212> DNA
<213> unknown
<308> M22760
<400> 171
TGTCATCCAG GAACTYAGAC CAACTTTAAA T

35
<210> 172
<211> 31
<212> DNA
<213> unknown
40
<308> M58028
<400> 172
TACCTTTAGC ATCTGYGACA CCTCCAACCTT C

45
<210> 173
<211> 31
<212> DNA
<213> unknown
<308> M58028
50
<400> 173
TTGGTGGCCT CACTGSCAGA ACCTGACTTT G

55

5
10
15
20
25
30
35
40
45
50
55

<210> 174
<211> 31
<212> DNA
<213> unknown
<308> M58028
<400> 174
TTCTCGCCCT GCCASCTGC ACATTGGCTT C

<210> 175
<211> 31
<212> DNA
<213> unknown
<308> M58028
<400> 175
ACAAGTGCCT CCAGCRCCAG AACCGTTATG A

<210> 176
<211> 31
<212> DNA
<213> unknown
<308> M58028
<400> 176
CTGACTGCGT GACCTSGGCC TGCCACCACT G

<210> 177
<211> 31
<212> DNA
<213> unknown
<308> M58028
<400> 177
AAGCTCAGGA GCGCCRTTCT GGTCTGGGCC C

<210> 178
<211> 31
<212> DNA
<213> unknown
<308> M58028
<400> 178
CCGAGCTGCT GTGGCYACAT TCCTGCAGTC T

<210> 179
<211> 31
<212> DNA
<213> unknown
<308> M58028
<400> 179
ACCAGGAGCT GCAGASCGCC AATGCCTCTG T

5
10
15
20
25
30
35
40
45
50
55

<210> 180
<211> 31
<212> DNA
<213> unknown
<308> M60784
<400> 180
CCCTTTCTAT GACAARCCTA TGCGTATCCA G

<210> 181
<211> 31
<212> DNA
<213> unknown
<308> U23852
<400> 181
TTGACTACCT GCGCASTGTG CTGGAGGACT T

<210> 182
<211> 31
<212> DNA
<213> unknown
<308> U23852
<400> 182
CTGCACATGT CTTGTRCATG TGTAGCCTGT G

<210> 183
<211> 31
<212> DNA
<213> unknown
<308> U23852
<400> 183
ATGTATGTCT TGGACRCTGT ACAAGGTACC C

<210> 184
<211> 31
<212> DNA
<213> unknown
<308> U30825
<400> 184
GACCGACGTG CGCGARAAGG ACTTGGAGGA C

<210> 185
<211> 31
<212> DNA
<213> unknown
<308> U30825
<400> 185
TGGAGGACCT GTTCTWCAAG TACGGCCGCA T

5
10
15
20
25
30
35
40
45
50
55

<210> 186
<211> 31
<212> DNA
<213> unknown
<308> U30825
<400> 186
ACCTGAAGGA TCACAYGCGA GAAGCTGGGG A

<210> 187
<211> 31
<212> DNA
<213> unknown
<308> U60975
<400> 187
GCCACAGTAA CAACCRACCC AATTTATACA T

<210> 188
<211> 31
<212> DNA
<213> unknown
<308> U60975
<400> 188
TGGCTCAGCG CCTCASTCAG CTCCTCAACC T

<210> 189
<211> 31
<212> DNA
<213> unknown
<308> U60975
<400> 189
TCTAGCAGTG CTGGARCCAG GTGGCGAGAG G

<210> 190
<211> 31
<212> DNA
<213> unknown
<308> U60975
<400> 190
CATGGGGAGA CCACGRCGGA ATCATCACGG C

<210> 191
<211> 31
<212> DNA
<213> unknown
<308> U60975
<400> 191
TTCTGGCCTG GAGACWGTAAG AAGCTTTGGC T

5
 <210> 192
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 192
 AATACCTGTG TCAAASAAGA GAACACCTGT C

10
 <210> 193
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 193
 CTGCAGGGAC TGGTCKGATG AAGCCAACTG T

15
 <210> 194
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 194
 GTGTGACGGG GATACRGACT GCCAGGATGG T

20
 <210> 195
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 195
 ACTGCCAGGA TGGTTSCGAT GAGGATCCAG T

25
 <210> 196
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 196
 TCAACTGTGA GAAGARGTGC AATGGATTCC G

30
 <210> 197
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 197
 ATTCCGCTGC CCAAAYGGCA CTTGCATCCC A

35
 <210> 198
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 198
 AATACCTGTG TCAAASAAGA GAACACCTGT C

40
 <210> 199
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 199
 AATACCTGTG TCAAASAAGA GAACACCTGT C

45
 <210> 200
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 200
 AATACCTGTG TCAAASAAGA GAACACCTGT C

50
 <210> 201
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 201
 AATACCTGTG TCAAASAAGA GAACACCTGT C

55
 <210> 202
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 202
 AATACCTGTG TCAAASAAGA GAACACCTGT C

EP 1 024 200 A2

5
 <210> 198
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 198
 TGATGAAGCC AACTGMGAAA ACCCCACAGA A

10
 <210> 199
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 199
 ACCGAGATTG TGCAGRTGGC TCTGACGAGG A

15
 <210> 200
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 200
 AAAAATGCCC TCTGCWTCTT GTGTATATAA T

20
 <210> 201
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 201
 CAATCTGACA GCTCAWACAT CCTATGAGAT T

25
 <210> 202
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 202
 CCTCTATCGC AACCCRAAGA GCTTGACTAC T

30
 <210> 203
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 203
 GACTTGTTGT ATGCARTTGC AGTCAAAGAT C

35
 <210> 204
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 204
 GACTTGTTGT ATGCARTTGC AGTCAAAGAT C

40
 <210> 205
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 205
 GACTTGTTGT ATGCARTTGC AGTCAAAGAT C

45
 <210> 206
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 206
 GACTTGTTGT ATGCARTTGC AGTCAAAGAT C

50
 <210> 207
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 207
 GACTTGTTGT ATGCARTTGC AGTCAAAGAT C

55
 <210> 208
 <211> 31
 <212> DNA
 <213> unknown
 <308> U60975
 <400> 208
 GACTTGTTGT ATGCARTTGC AGTCAAAGAT C

5
10
15
20
25
30
35
40
45
50
55

<210> 204
<211> 31
<212> DNA
<213> unknown
<308> U60975
<400> 204
CACCGGAGGC TGCAGRGCAG CTTACCGCC T

<210> 205
<211> 31
<212> DNA
<213> unknown
<308> X07743
<400> 205
AGCATGGAAC CAAAGYGGAT CAGAGAGGGC T

<210> 206
<211> 31
<212> DNA
<213> unknown
<308> X07743
<400> 206
GGTTCGGGAT ATCAAKAAGG CCATTAAATG C

<210> 207
<211> 31
<212> DNA
<213> unknown
<308> X07743
<400> 207
CATGATTGCT TCATCRCTGC TCAATGAGGG G

<210> 208
<211> 31
<212> DNA
<213> unknown
<308> X07743
<400> 208
CTGAAAACCC TTTCCCKGGAC AACCTGATG C

<210> 209
<211> 31
<212> DNA
<213> unknown
<308> X07743
<400> 209
GCACAGAGTG GATCARAGCC ATCCAGATGG C

5 <210> 210
 <211> 31
 <212> DNA
 <213> unknown
 <308> X07743
 <400> 210
 10 ACTGGGAAGT AAAGASACTC CTGCATTCCT C

 <210> 211
 <211> 31
 <212> DNA
 <213> unknown
 15 <308> X07743
 <400> 211
 CCTGAGGGAA GCCCAYGGAC AAGCTCAGTC C

20 <210> 212
 <211> 31
 <212> DNA
 <213> unknown
 <308> X07743
 <400> 212
 25 TTCCTGAGGT CCCCCYAGCT TAAAAAAAAA A

 <210> 213
 <211> 31
 <212> DNA
 30 <213> unknown
 <308> X07743
 <400> 213
 AAAAAAATCT GCCCCTGAT TCTAACACTC G

35 <210> 214
 <211> 31
 <212> DNA
 <213> unknown
 <308> X07743
 40 <400> 214
 ATCTTCTACT TGCTCYTGGC CTTGAGATCG T

 <210> 215
 <211> 31
 <212> DNA
 <213> unknown
 <308> X07743
 45 <400> 215
 50 CTGGTGAATG TGTCTYCGG CAGCTGCAGC A

55

5
10
<210> 216
<211> 31
<212> DNA
<213> unknown
<308> X07743
<400> 216
GGTGAATGTG TCTTGYGGCA GCTGCAGCAA G

15
20
<210> 217
<211> 31
<212> DNA
<213> unknown
<308> X07743
<400> 217
ACAGCACCCA TTGAARCAGA TATGTGTGTG A

25
30
<210> 218
<211> 31
<212> DNA
<213> unknown
<308> X07743
<400> 218
CCCATTGAAA CAGATRTGTG TGTGAAAGTA T

35
40
<210> 219
<211> 31
<212> DNA
<213> unknown
<308> X51521
<400> 219
CCTGAAGATG TGGCTSAGGA GCTCATCCAG G

45
50
<210> 220
<211> 31
<212> DNA
<213> unknown
<308> X51521
<400> 220
CCCCCCTGAG ACTGCGGTGC TCTTGGGGTC C

55
60
<210> 221
<211> 31
<212> DNA
<213> unknown
<308> X51521
<400> 221
GCAGGCCTGA TTCTCRGAT TATTCTCGAA T

5
<210> 222
<211> 31
<212> DNA
<213> unknown
<308> X51521
<400> 222
GGAGGACACT TGGATWTTTT TTTAGTTCTT T

10
<210> 223
<211> 31
<212> DNA
<213> unknown
15
<308> X51521
<400> 223
ATTAATGATC AGCTAYATAC TATTTATATA C

20
<210> 224
<211> 31
<212> DNA
<213> unknown
25
<308> X51521
<400> 224
CTATGCACAC TTTTAWTTTT GTCAGATTCA C

30
<210> 225
<211> 31
<212> DNA
<213> unknown
35
<308> X51521
<400> 225
GGTGGATTCC ATGCCRCAGA CATGAAATAA A

40
<210> 226
<211> 31
<212> DNA
<213> unknown
45
<308> X52882
<400> 226
ATTAAATACA CAGACRTAAG AGGCCAGCCA C

50
<210> 227
<211> 31
<212> DNA
<213> unknown
55
<308> X52882
<400> 227
GGTGGGATCC CAGGGYATGC CCAAGAGAAT C

5
5
10
15
20
25
30
35
40
45
50
55

<210> 228
<211> 31
<212> DNA
<213> unknown
<308> X52882
<400> 228
ATGATTTCAT GTGTGRTGAG ATGGAGCGCT C

<210> 229
<211> 31
<212> DNA
<213> unknown
<308> X52882
<400> 229
TTCTTGTTAT TCCARTACA CTAGCAGTTA A

<210> 230
<211> 31
<212> DNA
<213> unknown
<308> X52882
<400> 230
GAATTGATGA TCTTAYTAAA TTACATCCAG A

<210> 231
<211> 31
<212> DNA
<213> unknown
<308> AC002115
<400> 231
CAAGAACTAC AAGACYGCCC CTTTTGACAG C

<210> 232
<211> 31
<212> DNA
<213> unknown
<308> AC002115
<400> 232
TAATAAAAAC TCATTGAAA AGTGAGACTA T

<210> 233
<211> 31
<212> DNA
<213> unknown
<308> D13118
<400> 233
CACAGCAGCC AAGTTKATTG GTGCTGGGGC A

5
 <210> 234
 <211> 31
 <212> DNA
 <213> unknown
 <308> D13118
 <400> 234
 CATTGGCTAT GCCAGRAACC CGTCTCTCAA G

10
 <210> 235
 <211> 31
 <212> DNA
 <213> unknown
 <308> D13118
 <400> 235
 TTTGATGGTC GCCTTYCTCA TCCTCTTCGC C

15
 <210> 236
 <211> 31
 <212> DNA
 <213> unknown
 <308> D13118
 <400> 236
 TACTGCAACT CCACAYCATT CTTGGTGCTG G

20
 <210> 237
 <211> 31
 <212> DNA
 <213> unknown
 <308> D13118
 <400> 237
 AACTCCACAC CATTCYTGGT GCTGGGGTGT G

25
 <210> 238
 <211> 31
 <212> DNA
 <213> unknown
 <308> D13118
 <400> 238
 TGGTGCTGGG GTGTGYTAAG CTTTACCATT A

30
 <210> 239
 <211> 31
 <212> DNA
 <213> unknown
 <308> D26068
 <400> 239
 ACGGCTAGTC AGAGAMAAAG ACACAGATAA A

35
 <210> 234
 <211> 31
 <212> DNA
 <213> unknown
 <308> D13118
 <400> 234
 CATTGGCTAT GCCAGRAACC CGTCTCTCAA G

40
 <210> 235
 <211> 31
 <212> DNA
 <213> unknown
 <308> D13118
 <400> 235
 TTTGATGGTC GCCTTYCTCA TCCTCTTCGC C

45
 <210> 236
 <211> 31
 <212> DNA
 <213> unknown
 <308> D13118
 <400> 236
 TACTGCAACT CCACAYCATT CTTGGTGCTG G

50
 <210> 237
 <211> 31
 <212> DNA
 <213> unknown
 <308> D13118
 <400> 237
 AACTCCACAC CATTCYTGGT GCTGGGGTGT G

55
 <210> 238
 <211> 31
 <212> DNA
 <213> unknown
 <308> D13118
 <400> 238
 TGGTGCTGGG GTGTGYTAAG CTTTACCATT A

5
5
10
15
20
25
30
35
40
45
50
55

<210> 240
<211> 31
<212> DNA
<213> unknown
<308> D26068
<400> 240
TAAATTTAAA GGATTYTGCT ATGTAGAATT C

<210> 241
<211> 31
<212> DNA
<213> unknown
<308> D26068
<400> 241
AAGTCCCAGT CTGTCSTATG GGGAGCAGTT T

<210> 242
<211> 31
<212> DNA
<213> unknown
<308> D26068
<400> 242
AGTTTGGGGG CGGCCRCAG CAGGAGCCTG G

<210> 243
<211> 31
<212> DNA
<213> unknown
<308> D31890
<400> 243
CTGCTGCTGC CACCARCCAC ACCACTGATA A

<210> 244
<211> 31
<212> DNA
<213> unknown
<308> D31890
<400> 244
GGGATCACCT GACTGMCATC ACCTTAAAGG T

<210> 245
<211> 31
<212> DNA
<213> unknown
<308> D31890
<400> 245
ACCAAGAAGG GTGAGSTGAG CATCATTCGG T

5
<210> 246
<211> 31
<212> DNA
<213> unknown
<308> D31890
<400> 246
CTTTGTGAGG CAGAAMTTTA TCATCCGCTC T

10
<210> 247
<211> 31
<212> DNA
<213> unknown
15
<308> D31890
<400> 247
GGTTGGTGGC ATCGASCGGG TTTATGAAAT T

20
<210> 248
<211> 31
<212> DNA
<213> unknown
<308> D31890
25
<400> 248
GCACAACAGT TGGCASTTCT GTCTAGAAAA T

30
<210> 249
<211> 31
<212> DNA
<213> unknown
<308> HT3238
<400> 249
CAGAACGCTT CAGTTSTGCT CTGCAAGGAT A

35
<210> 250
<211> 31
<212> DNA
<213> unknown
40
<308> HT3238
<400> 250
CTCTGCAAGG ATATAWAATA ACTGATTGGT G

45
<210> 251
<211> 31
<212> DNA
<213> unknown
<308> HT3238
50
<400> 251
GCATTCTGCT TCAGASCAAG AATGCTGGGG C

55

5
10
15
20
25
30
35
40
45
50
55

<210> 252
<211> 31
<212> DNA
<213> unknown
<308> HT3238
<400> 252
ACCAGCCAGC TCCCGYTCGA ATCTGATGCT G

<210> 253
<211> 31
<212> DNA
<213> unknown
<308> HT3238
<400> 253
TGCTGTGGAA TGCTTWAATT ACCAACACTA T

<210> 254
<211> 31
<212> DNA
<213> unknown
<308> HT3238
<400> 254
GGAATTATTG GGGTCRAAGG TGCTAAAATC A

<210> 255
<211> 31
<212> DNA
<213> unknown
<308> HT3238
<400> 255
GGGTTGTAGA GTGCAWAAAG ATCATCCTTG A

<210> 256
<211> 31
<212> DNA
<213> unknown
<308> HT3238
<400> 256
AGAGCTCGGA ATCTTYCTCT TCCTCCACCA C

<210> 257
<211> 31
<212> DNA
<213> unknown
<308> HT3238
<400> 257
GTGATCTTGG TGGACYTATT ATTACTACAC A

EP 1 024 200 A2

5

15

20

30

40

45

55

5
<210> 264
<211> 31
<212> DNA
<213> unknown
<308> M20471
<400> 264
TCTTTTGGAC CAAACYTTTT TGTCTTTAGA G

10
<210> 265
<211> 31
<212> DNA
<213> unknown
15 <308> M22538
<400> 265
ATTTGCATAA GACAGYTATG CAAAATGGAG C

20
<210> 266
<211> 31
<212> DNA
<213> unknown
25 <308> M22538
<400> 266
CTTTTATAC AATGTWTAAT CGAAAGCCAG T

30
<210> 267
<211> 31
<212> DNA
<213> unknown
35 <308> M22538
<400> 267
GTGTACAAGC AGGCCKTTAA TTTATATTGA A

40
<210> 268
<211> 31
<212> DNA
<213> unknown
45 <308> U10439
<400> 268
GTGCTAATTG ACATGRAAAG GCAGGGGGAT G

50
<210> 269
<211> 31
<212> DNA
<213> unknown
55 <308> U10439
<400> 269
CAATCCCTGA GACCARAAGA AACGCAGAGT T

5
<210> 270
<211> 31
<212> DNA
<213> unknown
<308> U10439
<400> 270
AAGCCAAGGA CAGTGRAAAA TCAGAAGAAT C

10
<210> 271
<211> 31
<212> DNA
<213> unknown
15
<308> U10439
<400> 271
AGGGAATCGC TGTGTRAAAG GAGATTCTCT C

20
<210> 272
<211> 31
<212> DNA
<213> unknown
<308> U10439
25
<400> 272
CACCAAGGTG GAGAAYGGAG AAGGCACAAT C

30
<210> 273
<211> 31
<212> DNA
<213> unknown
<308> U10439
<400> 273
35
AACGTGCTGG GCCTGMAAGG GGCACCTGTTG A

40
<210> 274
<211> 31
<212> DNA
<213> unknown
<308> U10439
45
<400> 274
TGGCTGATGG CTATGRCCTG GAGATCCTGG A

50
<210> 275
<211> 31
<212> DNA
<213> unknown
<308> U10439
55
<400> 275
AAAAAGAACA TTTTTYTTCT ATTTAAGAAG C

5
10
15
20
25
30
35
40
45
50
55

<210> 276
<211> 31
<212> DNA
<213> unknown
<308> U10439
<400> 276
TTCTTTCCTT GTGATYTGAA TGTCTCCTTT T

<210> 277
<211> 31
<212> DNA
<213> unknown
<308> U10439
<400> 277
CCCTCAGAGG GCAAARAGGT GAACATAAAG G

<210> 278
<211> 31
<212> DNA
<213> unknown
<308> U10439
<400> 278
GCCTGCAGAT AATGCRYCAGC CATCCTCCCA T

<210> 279
<211> 31
<212> DNA
<213> unknown
<308> U10439
<400> 279
TCTCTGCTCA GATAAYCATG ACTTAGCAAG A

<210> 280
<211> 31
<212> DNA
<213> unknown
<308> U10439
<400> 280
CGTCACTGTT CCACCYGGTG TAATATCTCT C

<210> 281
<211> 31
<212> DNA
<213> unknown
<308> U10439
<400> 281
GCAGACCTGA TCTTTMTAGG GTTGACATAG A

5 <210> 282
 <211> 31
 <212> DNA
 <213> unknown
 <308> U10439
 <400> 282
 10 CTGTCAGTGC TTGAARCTGT TTCCTTTACT C
 <210> 283
 <211> 31
 <212> DNA
 <213> unknown
 15 <308> U10439
 <400> 283
 TTGGTGCCGT GGTGAWGGCT GCAGTCCAGT T
 <210> 284
 20 <211> 31
 <212> DNA
 <213> unknown
 <308> U10439
 25 <400> 284
 CTTTGCTTTT ATTCTRGTCT CTCTGTACTT T
 <210> 285
 <211> 31
 30 <212> DNA
 <213> unknown
 <308> U10439
 <400> 285
 TTGTCAGTGC TTGAARCTGT TTCCTTTACT C
 35 <210> 286
 <211> 31
 <212> DNA
 <213> unknown
 40 <308> U10439
 <400> 286
 CTGCAGTCCA GTTTTTRTGAT TCTGCTTTTA T
 <210> 287
 45 <211> 31
 <212> DNA
 <213> unknown
 <308> U57342
 50 <400> 287
 CACGTTCCGG GCCGCSACCC TGACCCTGCA G
 55

5
10
15
20
25
30
35
40
45
50
55

<210> 288
<211> 31
<212> DNA
<213> unknown
<308> U57342
<400> 288
AGGGCCCCGG GCCCTYAGCC TCTCTTGTAC A

<210> 289
<211> 31
<212> DNA
<213> unknown
<308> U57342
<400> 289
TTTCCCTCTCG ATTCCYATCC CCAATTTAAT A

<210> 290
<211> 31
<212> DNA
<213> unknown
<308> U57342
<400> 290
TACCATTGAT GACTTYGCCT CTGCATCTAC T

<210> 291
<211> 31
<212> DNA
<213> unknown
<308> U57342
<400> 291
GCCCCCTCCA TCCCTYAGCT ACCCTGGATC T

<210> 292
<211> 31
<212> DNA
<213> unknown
<308> U57342
<400> 292
GGAGCAGACT GGTTTCYTAAC TCTCTCCCTC C

<210> 293
<211> 31
<212> DNA
<213> unknown
<308> U62962
<400> 293
AATTGGGATG CAGCCRTGGA AGACCTTACA C

5
10
15
20
25
30
35
40
45
50
55

<210> 294
<211> 31
<212> DNA
<213> unknown
<308> U62962
<400> 294
GGGCAACTCA AGATTYTGGC TTCTACTGAA G

<210> 295
<211> 31
<212> DNA
<213> unknown
<308> X13238
<400> 295
TTTCTACAGA AACTAYGATG TCATGAAAGA T

<210> 296
<211> 31
<212> DNA
<213> unknown
<308> X14046
<400> 296
CATCTGGATC CTCATYGACA AGACCAGCTT C

<210> 297
<211> 31
<212> DNA
<213> unknown
<308> X14046
<400> 297
ACAGATCACC CTGGGMATCC TCATCTCCAC T

<210> 298
<211> 31
<212> DNA
<213> unknown
<308> X14046
<400> 298
TGGCACTACC CGCAGSACTG GTTCCAAGTC C

<210> 299
<211> 31
<212> DNA
<213> unknown
<308> X14046
<400> 299
CCCGCAGGAC TGGTTYCAAG TCCTCATCCT G

EP 1 024 200 A2

5
 <210> 300
 <211> 31
 <212> DNA
 <213> unknown
 <308> X14046
 <400> 300
 TCACGTGCGC TGGGCWCTTC CCTGCTGCCT G

10

15
 <210> 301
 <211> 31
 <212> DNA
 <213> unknown
 <308> X57346
 <400> 301
 AGGAAGGAAG AGGTCMTCTC GCTCGGAGCT T

20

25
 <210> 302
 <211> 31
 <212> DNA
 <213> unknown
 <308> X57346
 <400> 302
 TGCAGCCCTC CCACGSCAGA GTCTCCAGAG A

30

35
 <210> 303
 <211> 31
 <212> DNA
 <213> unknown
 <308> X57346
 <400> 303
 TGAGCAGGCT GAGCGMTATG ATGATATGGC T

40

45
 <210> 304
 <211> 31
 <212> DNA
 <213> unknown
 <308> X57346
 <400> 304
 GCATGAACTC TCCAASGAAG AGAGAAATCT G

50

55
 <210> 305
 <211> 31
 <212> DNA
 <213> unknown
 <308> X57346
 <400> 305
 TCGTGCTTTG TGATCYGTCC AGTGTCACCTC T

EP 1 024 200 A2

5
 <210> 306
 <211> 31
 <212> DNA
 <213> unknown
 <308> X74008
 <400> 306
 CTCCAAGGGG TATGAYCACA AAGCAAGCAA A

10
 <210> 307
 <211> 31
 <212> DNA
 <213> unknown
 <308> X74008
 <400> 307
 AGGTGGTATG GTTGTRTACA CATGCTGCCT T

15
 <210> 308
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z29505
 <400> 308
 ACAGCTCCAT GACCAWCAGT ACCGCGGCCA G

20
 <210> 309
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z29505
 <400> 309
 ATCAAAGAGA TCCGCSAGAG TACGGGGGCG C

25
 <210> 310
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z37166
 <400> 310
 TGGGCAGGTG TCTGTRCTGG TGATGTGTCA C

30
 <210> 311
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z37166
 <400> 311
 TTTCAGATCA GCAAGRAATA TGAGCGCTTC T

35
 <210> 310
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z37166
 <400> 310
 TGGGCAGGTG TCTGTRCTGG TGATGTGTCA C

40
 <210> 311
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z37166
 <400> 311
 TTTCAGATCA GCAAGRAATA TGAGCGCTTC T

45
 <210> 311
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z37166
 <400> 311
 TTTCAGATCA GCAAGRAATA TGAGCGCTTC T

50
 <210> 311
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z37166
 <400> 311
 TTTCAGATCA GCAAGRAATA TGAGCGCTTC T

55

EP 1 024 200 A2

5
 <210> 312
 <211> 31
 <212> DNA
 <213> unknown
 <308> AC002045
 <400> 312
 TCTCCCTCTG CACAASTTAC CTTTGTGTGT C

10
 <210> 313
 <211> 31
 <212> DNA
 <213> unknown
 <308> AC002045
 <400> 313
 GGTGGAAACC AAAGTWCAG CTAAAATCTG T

15
 <210> 314
 <211> 31
 <212> DNA
 <213> unknown
 <308> AC002045
 <400> 314
 GTGGAAACCA AAGTTMGAGC TAAAATCTGT A

20
 <210> 315
 <211> 31
 <212> DNA
 <213> unknown
 <308> D26599
 <400> 315
 CACGTGAGAG GGCAGWGGAA CTCCTTAGGA A

25
 <210> 316
 <211> 31
 <212> DNA
 <213> unknown
 <308> D30655
 <400> 316
 GTTCTCTATC ATTTARTAAT ATACTTGTGG A

30
 <210> 317
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08096
 <400> 317
 CCATTGGTCG CGGGCYTGGT GATCTGCCTC G

35
 <210> 318
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08096
 <400> 318
 CCATTGGTCG CGGGCYTGGT GATCTGCCTC G

40
 <210> 319
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08096
 <400> 319
 CCATTGGTCG CGGGCYTGGT GATCTGCCTC G

45
 <210> 320
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08096
 <400> 320
 CCATTGGTCG CGGGCYTGGT GATCTGCCTC G

50
 <210> 321
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08096
 <400> 321
 CCATTGGTCG CGGGCYTGGT GATCTGCCTC G

55
 <210> 322
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08096
 <400> 322
 CCATTGGTCG CGGGCYTGGT GATCTGCCTC G

EP 1 024 200 A2

5
 <210> 318
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08096
 <400> 318
 10 AGCTGAATCA CACAGSACCT CAGCAGGACC C

15
 <210> 319
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08096
 <400> 319
 CCTGCATGGA CCAGARCTGG ACAAGGGGCA G

20
 <210> 320
 <211> 31
 <212> DNA
 <213> unknown
 <08> L08096

25
 <400> 320
 CTGCATGGAC CAGAGMTGGA CAAGGGGCAG C

30
 <210> 321
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08096
 <400> 321

35
 TGCATGGACC AGAGCWGGAC AAGGGGCAGC T

40
 <210> 322
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08096
 <400> 322
 TCTACATGGT ACACAKCCAG GTGACGCTGG C

45
 <210> 323
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08096

50
 <400> 323
 CACATCCAGG TGACGYTGGC CATCTGCTCC T

55

5
 <210> 324
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08096
 <400> 324
 GACGCTGGCC ATCTGYTCCT CCACGACGGC C

10
 <210> 325
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08096
 <400> 325
 CCCCACCACC CTGGCYGTGG GAATCTGCTC T

15
 <210> 326
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08096
 <400> 326
 CCGTAGCATC AGCCTSCTGC GTCTCAGCTT C

20
 <210> 327
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08096
 <400> 327
 TAGCATCAGC CTGCTSCGTC TCAGCTTCCA C

25
 <210> 328
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 328
 GAATGCCGGC CAAGCRACAG GCTGTCACCA A

30
 <210> 329
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 329
 TCATTGGCCG GCGATRTGAT GATCCTGAAG T

35
 <210> 329
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 329
 TCATTGGCCG GCGATRTGAT GATCCTGAAG T

40
 <210> 329
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 329
 TCATTGGCCG GCGATRTGAT GATCCTGAAG T

45
 <210> 329
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 329
 TCATTGGCCG GCGATRTGAT GATCCTGAAG T

50
 <210> 329
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 329
 TCATTGGCCG GCGATRTGAT GATCCTGAAG T

55

5
 <210> 330
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 330
 GCATTTGTGT TGATGRAGAT GAAAGAGACT G

10
 <210> 331
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 331
 TCAATGACTC GCAGASACAG GCCACTAAAG A

15
 <210> 332
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 332
 TATCCTGGAA ATTCARAAAG GAGTATTTGA G

20
 <210> 333
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 333
 TGCTGAAAAG GCTAARTGTG AACTCTCCTC A

25
 <210> 334
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 334
 TTGTCACTGA TCTAAYCAGA AGGACTATCG C

30
 <210> 335
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 335
 TGTGTTGGCC GGCGAKGTCA CGGATGTGCT G

35
 <210> 336
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 336
 TGTGTTGGCC GGCGAKGTCA CGGATGTGCT G

40
 <210> 337
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 337
 TGTGTTGGCC GGCGAKGTCA CGGATGTGCT G

45
 <210> 338
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 338
 TGTGTTGGCC GGCGAKGTCA CGGATGTGCT G

50
 <210> 339
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 339
 TGTGTTGGCC GGCGAKGTCA CGGATGTGCT G

55
 <210> 340
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 340
 TGTGTTGGCC GGCGAKGTCA CGGATGTGCT G

5
 <210> 336
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 336
 ACTTATTAAT AGGAAYACCA CTATTCCAAC C

10
 <210> 337
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 337
 GGTATTCTCT ACTGCGCTG ATGGTCAAAC G

15
 <210> 338
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 338
 CTTCAGCAGG CATCAYTGAA GCTGTTCGAA A

20
 <210> 339
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 339
 TATAAAGATT TAACTRTTTT ATGCTGAAGT G

25
 <210> 340
 <211> 31
 <212> DNA
 <213> unknown
 <308> L11066
 <400> 340
 GCTACCTTCT GCCCTSTGTC TGGCACCTAC A

30
 <210> 341
 <211> 31
 <212> DNA
 <213> unknown
 <308> U05259
 <400> 341
 GAAACGATGG CAGAASGAGA AGCTCGGGTT G

35
 <210> 341
 <211> 31
 <212> DNA
 <213> unknown
 <308> U05259
 <400> 341
 GAAACGATGG CAGAASGAGA AGCTCGGGTT G

40
 <210> 341
 <211> 31
 <212> DNA
 <213> unknown
 <308> U05259
 <400> 341
 GAAACGATGG CAGAASGAGA AGCTCGGGTT G

45
 <210> 341
 <211> 31
 <212> DNA
 <213> unknown
 <308> U05259
 <400> 341
 GAAACGATGG CAGAASGAGA AGCTCGGGTT G

50
 <210> 341
 <211> 31
 <212> DNA
 <213> unknown
 <308> U05259
 <400> 341
 GAAACGATGG CAGAASGAGA AGCTCGGGTT G

55

EP 1 024 200 A2

5
 <210> 342
 <211> 31
 <212> DNA
 <213> unknown
 <308> U05259
 <400> 342
 ACCTGGACGA CTGCTYCATG TATGAGGACA T

10

15
 <210> 343
 <211> 31
 <212> DNA
 <213> unknown
 <308> U05259
 <400> 343
 CCTCAACATA GGAGAWGTCC AGCTGGAGAA G

20

25
 <210> 344
 <211> 31
 <212> DNA
 <213> unknown
 <308> U44839
 <400> 344
 GGGCCCAGCT CTGGARTCAC GAACAGGTGC C

30

35
 <210> 345
 <211> 31
 <212> DNA
 <213> unknown
 <308> U44839
 <400> 345
 GAAGCATGAC TGCCTYGGGT ACGTGATGAA G

40

45
 <210> 346
 <211> 31
 <212> DNA
 <213> unknown
 <308> U44839
 <400> 346
 TGATGACAAC AGCGTMTCCC CTGTCAATGA G

50

55
 <210> 347
 <211> 31
 <212> DNA
 <213> unknown
 <308> U44839
 <400> 347
 ACGCCAGGAC GTGGCRCGAC GCCTGCTGTC C

EP 1 024 200 A2

5
 <210> 348
 <211> 31
 <212> DNA
 <213> unknown
 <308> U46025
 <400> 348
 AATAATGCAA AATAACKGACC CTCACTCCCA A

10
 <210> 349
 <211> 31
 <212> DNA
 <213> unknown
 <308> U46025
 <400> 349
 AGGTCTGCCG CATCTWCCTG CTGCGCATCC T

15
 <210> 350
 <211> 31
 <212> DNA
 <213> unknown
 <308> U46025
 <400> 350
 TGAGGGCTCC TCAASTCTG AGCAAGACCA G

20
 <210> 351
 <211> 31
 <212> DNA
 <213> unknown
 <308> U46025
 <400> 351
 AGGGCGAGGA CTCGGYTGTG TTGATGGAGA G

25
 <210> 352
 <211> 31
 <212> DNA
 <213> unknown
 <308> U46025
 <400> 352
 GAGGACTCGG CTGTGWTGAT GGAGAGACTG T

30
 <210> 353
 <211> 31
 <212> DNA
 <213> unknown
 <308> U46025
 <400> 353
 CCTCTGCCAC ATCTAYCACC ATGCTCTGCA C

35
 <210> 352
 <211> 31
 <212> DNA
 <213> unknown
 <308> U46025
 <400> 352
 GAGGACTCGG CTGTGWTGAT GGAGAGACTG T

40
 <210> 353
 <211> 31
 <212> DNA
 <213> unknown
 <308> U46025
 <400> 353
 CCTCTGCCAC ATCTAYCACC ATGCTCTGCA C

45
 <210> 353
 <211> 31
 <212> DNA
 <213> unknown
 <308> U46025
 <400> 353
 CCTCTGCCAC ATCTAYCACC ATGCTCTGCA C

50
 <210> 353
 <211> 31
 <212> DNA
 <213> unknown
 <308> U46025
 <400> 353
 CCTCTGCCAC ATCTAYCACC ATGCTCTGCA C

55

5
<210> 354
<211> 31
<212> DNA
<213> unknown
<308> U46025
<400> 354
TATCATCAAT GAGAARATGA ATGGGAAAGT G

10
<210> 355
<211> 31
<212> DNA
<213> unknown
<308> U46025
15
<400> 355
ACTGCCCAGC AGAACMTGGC TCTGCAGCTG G

20
<210> 356
<211> 31
<212> DNA
<213> unknown
<308> U70439
25
<400> 356
GCTAGCAAAC CCTTCYGACG GCCCTCGCTG C

30
<210> 357
<211> 31
<212> DNA
<213> unknown
<308> U70439
35
<400> 357
CTGGAGCTGA GGAACYGGAC CCCGGCAGCT G

40
<210> 358
<211> 31
<212> DNA
<213> unknown
<308> U70439
45
<400> 358
GAACTTAGAG TTCCTYAGTT TAATAAATGT A

50
<210> 359
<211> 31
<212> DNA
<213> unknown
<308> U70439
55
<400> 359
CTCAGTTTCA AATCTYCCCA AGCTGCCTAA A

5
5
10
15
20
25
30
35
40
45
50
55

<210> 360
<211> 31
<212> DNA
<213> unknown
<308> U70439
<400> 360
CATGTTAGCT GAAAARCTTC CAAATCTCAC A

<210> 361
<211> 31
<212> DNA
<213> unknown
<308> U70439
<400> 361
TTCCAAATCT CACACRTCTA AACTTAAGTG G

<210> 362
<211> 31
<212> DNA
<213> unknown
<308> U70439
<400> 362
GGAACCTTTG AAAAARTTAG AATGTCTGAA A

<210> 363
<211> 31
<212> DNA
<213> unknown
<308> U70439
<400> 363
CCCCAGCTTA CCTACWTGGA TGGCTATGAC C

<210> 364
<211> 31
<212> DNA
<213> unknown
<308> U70439
<400> 364
TGGATGGCTA TGACCRAGAG GACCAGGAAG C

<210> 365
<211> 31
<212> DNA
<213> unknown
<308> X07203
<400> 365
GGGTCTTCTG ATGATYCCAG CAGGGATCTA T

5
 <210> 366
 <211> 31
 <212> DNA
 <213> unknown
 <308> X07203
 <400> 366
 TCATTTTCTC CATCARCAAC CAGGGAGACT G

10
 <210> 367
 <211> 31
 <212> DNA
 <213> unknown
 <308> X15822
 <400> 367
 GCAGGAGTGA CTTTCARTCAT CCCAGCAATC G

15
 <210> 368
 <211> 31
 <212> DNA
 <213> unknown
 <308> X16663
 <400> 368
 GTACGGAGTT GAGAGRGACA GGGCAGACAA G

20
 <210> 369
 <211> 31
 <212> DNA
 <213> unknown
 <308> X16663
 <400> 369
 AGAAGTGGAG AAGCAYACAT CTCAGAAAGA T

25
 <210> 370
 <211> 31
 <212> DNA
 <213> unknown
 <308> X16663
 <400> 370
 GACGCCATA GAAGCYGCTT CTAGTGGTGC C

30
 <210> 371
 <211> 31
 <212> DNA
 <213> unknown
 <308> X16663
 <400> 371
 GCCGCTTCTA GTGGTRCCCG TGGGCTGAAG G

35
 <210> 372
 <211> 31
 <212> DNA
 <213> unknown
 <308> X16663
 <400> 372
 GACGCCATA GAAGCYGCTT CTAGTGGTGC C

40
 <210> 373
 <211> 31
 <212> DNA
 <213> unknown
 <308> X16663
 <400> 373
 GACGCCATA GAAGCYGCTT CTAGTGGTGC C

45
 <210> 374
 <211> 31
 <212> DNA
 <213> unknown
 <308> X16663
 <400> 374
 GACGCCATA GAAGCYGCTT CTAGTGGTGC C

50
 <210> 375
 <211> 31
 <212> DNA
 <213> unknown
 <308> X16663
 <400> 375
 GACGCCATA GAAGCYGCTT CTAGTGGTGC C

55
 <210> 376
 <211> 31
 <212> DNA
 <213> unknown
 <308> X16663
 <400> 376
 GACGCCATA GAAGCYGCTT CTAGTGGTGC C

5
10
15
20
25
30
35
40
45
50
55

<210> 372
<211> 31
<212> DNA
<213> unknown
<308> X16663
<400> 372
GAAGAGCCAG TGTACRAAGC AGAGCCTGAG C

<210> 373
<211> 31
<212> DNA
<213> unknown
<308> X16663
<400> 373
GAGATGGACA GGCATRAGCA GGAGGATGAA C

<210> 374
<211> 31
<212> DNA
<213> unknown
<308> X16663
<400> 374
GATCATCAGG CTGCCYGGCT GGGGCTGGGG C

<210> 375
<211> 31
<212> DNA
<213> unknown
<308> X16663
<400> 375
ATCTCAGCTG TGGCTSTATA TGATTACCAA G

<210> 376
<211> 31
<212> DNA
<213> unknown
<308> X16663
<400> 376
CTTCCCTGCA AATTAYGTCA AGCTTCTGGA G

<210> 377
<211> 31
<212> DNA
<213> unknown
<308> X16663
<400> 377
ACTGTGATTT CCCATKTCCA AAGTGGCTCT G

EP 1 024 200 A2

5
 <210> 378
 <211> 31
 <212> DNA
 <213> unknown
 <308> X16663
 <400> 378
 CCCTCCCTAT TCCTGMTGCA AATGTCTAAC C

10
 <210> 379
 <211> 31
 <212> DNA
 <213> unknown
 <308> X16663
 <400> 379
 GTCTAACCAG ATGAGKTTCT GGACAGACTT C

15
 <210> 380
 <211> 31
 <212> DNA
 <213> unknown
 <308> X16663
 <400> 380
 CTCCTTCCCA ATGAAYACCT CTCTGCCACC C

20
 <210> 381
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 381
 GAAGACTGAG CTGATSCAGA AGGCCAAGCT G

25
 <210> 382
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 382
 AGACTGAGCT GATCCWGAAG GCCAAGCTGG C

30
 <210> 383
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 383
 GGCCGAGCGC TACGAYGACA TGGCCACCTG C

35
 <210> 384
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 384
 GCGGAGGCGC TACGAYGACA TGGCCACCTG C

40
 <210> 385
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 385
 GCGGAGGCGC TACGAYGACA TGGCCACCTG C

45
 <210> 386
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 386
 GCGGAGGCGC TACGAYGACA TGGCCACCTG C

50
 <210> 387
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 387
 GCGGAGGCGC TACGAYGACA TGGCCACCTG C

55
 <210> 388
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 388
 GCGGAGGCGC TACGAYGACA TGGCCACCTG C

EP 1 024 200 A2

5
 <210> 384
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 384
 GGTCATCTCT AGCATYGAGC AGAAGACCGA C

10
 <210> 385
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 385
 CATCGAGCAG AAGACYGACA CCTCCGACAA G

20
 <210> 386
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 386
 GGCTTTTGAT GAGGCSATTG CTGAACTGA T

30
 <210> 387
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 387
 CTGAATGAAG ACTCAKACAA AGACAGCACC C

35
 <210> 388
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 388
 CACCCTCATC ATGCARTTGC TTAGAGACAA C

45
 <210> 389
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 389
 GGCTGAAAAC TAAATMCATA CAGGGTGTCA T

55

EP 1 024 200 A2

5
 <210> 390
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 390
 TCCATACAGG GTGTCMTCCT TCTTTCCTTC A

10
 <210> 391
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 391
 TTTAAGTAAT CTGAAMCAGT TCTGCAAGTG A

15
 <210> 392
 <211> 31
 <212> DNA
 <213> unknown
 <308> X56468
 <400> 392
 AAATTAAAGT ACACARGTCC AAGTCTAAAA C

20
 <210> 393
 <211> 31
 <212> DNA
 <213> unknown
 <308> X57351
 <400> 393
 AGGACCTCCA CACCAYTAAC AAGATGAGCC T

25
 <210> 394
 <211> 31
 <212> DNA
 <213> unknown
 <308> X57351
 <400> 394
 TGCATTTGAC AAATGSCAGG AAGAGGAAAC T

30
 <210> 395
 <211> 31
 <212> DNA
 <213> unknown
 <308> X57351
 <400> 395
 TCTCCTGTCA ACAGCRGCCA GCCTCCCAAC T

35
 <210> 396
 <211> 31
 <212> DNA
 <213> unknown
 <308> X57351
 <400> 396
 TCTCCTGTCA ACAGCRGCCA GCCTCCCAAC T

40
 <210> 397
 <211> 31
 <212> DNA
 <213> unknown
 <308> X57351
 <400> 397
 TCTCCTGTCA ACAGCRGCCA GCCTCCCAAC T

45
 <210> 398
 <211> 31
 <212> DNA
 <213> unknown
 <308> X57351
 <400> 398
 TCTCCTGTCA ACAGCRGCCA GCCTCCCAAC T

50
 <210> 399
 <211> 31
 <212> DNA
 <213> unknown
 <308> X57351
 <400> 399
 TCTCCTGTCA ACAGCRGCCA GCCTCCCAAC T

55
 <210> 400
 <211> 31
 <212> DNA
 <213> unknown
 <308> X57351
 <400> 400
 TCTCCTGTCA ACAGCRGCCA GCCTCCCAAC T

5
<210> 396
<211> 31
<212> DNA
<213> unknown
<308> X57351
<400> 396
CTTCATAGCA TTCGCSTACT CCGTGAAGTC T

10
<210> 397
<211> 31
<212> DNA
<213> unknown
15
<308> X57351
<400> 397
CTTCATGACC ATTCTSCTCG TCATCATCCC A

20
<210> 398
<211> 31
<212> DNA
<213> unknown
<308> X57351
<400> 398
25
ATGACCATTC TGCTCRTCAT CATCCCAGTG T

30
<210> 399
<211> 31
<212> DNA
<213> unknown
<308> X59892
<400> 399
35
AGTGAGCCCCG CATCTYTGCT GGAGCTGTTC A

40
<210> 400
<211> 31
<212> DNA
<213> unknown
<308> X59892
<400> 400
45
AGCTACAAAG CTGCCCKCGG GGAGGATTAC A

50
<210> 401
<211> 31
<212> DNA
<213> unknown
<308> X59892
<400> 401
55
CCTAAACCAG CCCTGYTGCA CTCCACCTTC T

EP 1 024 200 A2

5
 <210> 402
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59892
 <400> 402
 AATTGATAAG TAAAAKCTTC GTTATACATT T

10
 <210> 403
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59892
 <400> 403
 AGACTTGGGG GTGAARCTTT CCAGTTTACT G

15
 <210> 404
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 404
 TCCGGTACAG AATGTWTTGC CAAGTACAAC T

20
 <210> 405
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 405
 CTTCCACGGC ACTGCGYAGC AGGACCTGCC C

25
 <210> 406
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 406
 GGTGCAGCTC CTGGGYGTGA TCGTGGAGGA G

30
 <210> 407
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 407
 GACTACCTGC GGTCTMGGGG TCGGTCAGTG C

35
 <210> 408
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 408
 GGTGCTGCTC CTGGGYGTGA TCGTGGAGGA G

40
 <210> 409
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 409
 GGTGCTGCTC CTGGGYGTGA TCGTGGAGGA G

45
 <210> 410
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 410
 GGTGCTGCTC CTGGGYGTGA TCGTGGAGGA G

50
 <210> 411
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 411
 GGTGCTGCTC CTGGGYGTGA TCGTGGAGGA G

55
 <210> 412
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 412
 GGTGCTGCTC CTGGGYGTGA TCGTGGAGGA G

EP 1 024 200 A2

5
 <210> 408
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 408
 TGCTGGCACC TGGACRCCGC CATGCGGCCC T

10
 <210> 409
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 409
 TGCCCCCTGCT CACTGRGCCC GAGCCTGAAC T

15
 <210> 410
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 410
 GAGGAAGGAG GCCACRGAGC GGGAGGCAGC G

20
 <210> 411
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 411
 ATTCCTTTCC TTTTTKGAGA TTTTTTTTCC G

25
 <210> 412
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 412
 CTTGTGAGAT GGAATYGTA TAAACCACGC C

30
 <210> 413
 <211> 31
 <212> DNA
 <213> unknown
 <308> Z47055
 <400> 413
 TGGAGTACAA TGCCAYTGGA GGCAAGTATA A

35
 <210> 414
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 414
 TTTTCTTTCC TTTTCTGAGA TTTTCTTTCC G

40
 <210> 415
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 415
 TTTTCTTTCC TTTTCTGAGA TTTTCTTTCC G

45
 <210> 416
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 416
 TTTTCTTTCC TTTTCTGAGA TTTTCTTTCC G

50
 <210> 417
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 417
 TTTTCTTTCC TTTTCTGAGA TTTTCTTTCC G

55
 <210> 418
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59932
 <400> 418
 TTTTCTTTCC TTTTCTGAGA TTTTCTTTCC G

5
 <210> 414
 <211> 31
 <212> DNA
 <213> unknown
 <308> STSG-13126
 <400> 414
 10 TCCGGCCCTA TCGTCYTTCC TGTGGGCCAT T

15
 <210> 415
 <211> 31
 <212> DNA
 <213> unknown
 <308> STSG-13353
 <400> 415
 GAAATAGGAC CAATCYGAGG CTTTCGTGAG A

20
 <210> 416
 <211> 31
 <212> DNA
 <213> unknown
 <308> STSG-13369
 25 <400> 416
 GTGCTTTGGT ATATASAGGT AACAAATGTAC T

30
 <210> 417
 <211> 31
 <212> DNA
 <213> unknown
 <308> STSG-13448
 <400> 417
 ATGACAGATT CAAGAYGATG TCGCCTCCTG T

35
 <210> 418
 <211> 31
 <212> DNA
 <213> unknown
 40 <308> D00761
 <400> 418
 GCCATGTATT CGGCTSCTGG CAGAGACTTG G

45
 <210> 419
 <211> 31
 <212> DNA
 <213> unknown
 <308> D00761
 50 <400> 419
 TCAAGGCTGG AGGCTSAGCA AGTGCCATGC T

55

5
<210> 420
<211> 31
<212> DNA
<213> unknown
<308> D28137
<400> 420
GGCCCAGAAG GGCTTYCAGG ATGTGGAGGC C

10
<210> 421
<211> 31
<212> DNA
<213> unknown
<308> D43950
<400> 421
TTGCCAAGCT GATGGYGGAA CTGTCCAAGT C

15
<210> 422
<211> 31
<212> DNA
<213> unknown
<308> D43950
<400> 422
CTCGTGTTC TATTGWACAC CTGGACAAGA T

20
<210> 423
<211> 31
<212> DNA
<213> unknown
<308> D43950
<400> 423
TCTTGTACAG GAGATSTCAT TTGGGACAAC T

25
<210> 424
<211> 31
<212> DNA
<213> unknown
<308> D43950
<400> 424
TAATCGTGGG GGTACYATCT CAACTGCTTT T

30
<210> 425
<211> 31
<212> DNA
<213> unknown
<308> HT3200
<400> 425
ACGCCATCCG CGAGGYCCTG GAGCTGGACT C

35
40
45
50
55

EP 1 024 200 A2

5
 <210> 426
 <211> 31
 <212> DNA
 <213> unknown
 <308> J03077
 <400> 426
 10 TCGCGAGTCA GACGGYGCTA TGTACGCCCT C

15
 <210> 427
 <211> 31
 <212> DNA
 <213> unknown
 <308> J03077
 <400> 427
 20 ATAAATATGG ATGGCRAGCT CCTAGGCCTC T

25
 <210> 428
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08666
 <400> 428
 AAAGTTACTG GGACCYTGGA GACCAAATAC A

30
 <210> 429
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08666
 <400> 429
 35 AGGCTTTTTT CCCCCMAGAA GATGATCAAA A

40
 <210> 430
 <211> 31
 <212> DNA
 <213> unknown
 <308> L08666
 <400> 430
 TGCAGTCACC TATACRTTAT TTAAATGTAT T

45
 <210> 431
 <211> 31
 <212> DNA
 <213> unknown
 <308> L22009
 50 <400> 431
 AACTCATTAT GATCCRCCAC GAAAGCTTAT G

55

5
 <210> 432
 <211> 31
 <212> DNA
 <213> unknown
 <308> L22009
 <400> 432
 TGATTTATGC TTACTYTAAG TGGAAATCAG G

10
 <210> 433
 <211> 31
 <212> DNA
 <213> unknown
 <308> L22009
 <400> 433
 CTCAAGTTAG TTATAKGATG TACTTAAGCA G

15
 <210> 434
 <211> 31
 <212> DNA
 <213> unknown
 <308> L22009
 <400> 434
 TTAGTTATAG GATGTWCTTA AGCAGTAAGC G

20
 <210> 435
 <211> 31
 <212> DNA
 <213> unknown
 <308> L22009
 <400> 435
 CTTTTTTTTT TTTTAMGTTA ATGGGAAAAA T

25
 <210> 436
 <211> 31
 <212> DNA
 <213> unknown
 <308> L22009
 <400> 436
 TTGGTGTTTG TATGARATTC TGAGGCCTTG A

30
 <210> 437
 <211> 31
 <212> DNA
 <213> unknown
 <308> L28010
 <400> 437
 TTCTCATGCA AAATTYTCTT CTAGCATGTG A

35
 <210> 438
 <211> 31
 <212> DNA
 <213> unknown
 <308> L28010
 <400> 438
 TTCTCATGCA AAATTYTCTT CTAGCATGTG A

40
 <210> 439
 <211> 31
 <212> DNA
 <213> unknown
 <308> L28010
 <400> 439
 TTCTCATGCA AAATTYTCTT CTAGCATGTG A

45
 <210> 440
 <211> 31
 <212> DNA
 <213> unknown
 <308> L28010
 <400> 440
 TTCTCATGCA AAATTYTCTT CTAGCATGTG A

50
 <210> 441
 <211> 31
 <212> DNA
 <213> unknown
 <308> L28010
 <400> 441
 TTCTCATGCA AAATTYTCTT CTAGCATGTG A

55
 <210> 442
 <211> 31
 <212> DNA
 <213> unknown
 <308> L28010
 <400> 442
 TTCTCATGCA AAATTYTCTT CTAGCATGTG A

5
10
15
20
25
30
35
40
45
50
55

<210> 438
<211> 31
<212> DNA
<213> unknown
<308> L38951
<400> 438
TGCCAAGGAT TGTTAYCCTG CTGTCCAGAA A

<210> 439
<211> 31
<212> DNA
<213> unknown
<308> L38951
<400> 439
TTCTGTCTTT TGGCCRGTC CGAGTGGAAT G

<210> 440
<211> 31
<212> DNA
<213> unknown
<308> L38951
<400> 440
GTAAAAGGCA GCCTGRCCTT TGCTACTGAG G

<210> 441
<211> 31
<212> DNA
<213> unknown
<308> L38951
<400> 441
TGTACCTTTG CCACAYGGTA CTGTATGCTT G

<210> 442
<211> 31
<212> DNA
<213> unknown
<308> L38951
<400> 442
CATGGTACTG TATGCTGCC AGCTAGAAGG A

<210> 443
<211> 31
<212> DNA
<213> unknown
<308> L38951
<400> 443
TACTGTATGC TTGCCRGCTA GAAGGAGGGT C

5
 <210> 444
 <211> 31
 <212> DNA
 <213> unknown
 <308> L38951
 <400> 444
 TATCCACATT CTCAASTTCA AGTCATTGCT G

10
 <210> 445
 <211> 31
 <212> DNA
 <213> unknown
 <308> L38951
 <400> 445
 AAGTTCTGTC TACTGRTGCA GCACAAGAGA T

15
 <210> 446
 <211> 31
 <212> DNA
 <213> unknown
 <308> L38951
 <400> 446
 CAATTGTTAC AAATARCGCA GACTTCAAAA A

20
 <210> 447
 <211> 31
 <212> DNA
 <213> unknown
 <308> L38951
 <400> 447
 AAACAAACCA AAATTKAAAT GATCAGAATT G

25
 <210> 448
 <211> 31
 <212> DNA
 <213> unknown
 <308> L38951
 <400> 448
 CCAGAAAATT GTGCCRAAGA GTTTAGAAAA A

30
 <210> 449
 <211> 31
 <212> DNA
 <213> unknown
 <308> L38951
 <400> 449
 AAGAGTTTAG AAAAYAAAT ATACAATAAA A

35
 <210> 448
 <211> 31
 <212> DNA
 <213> unknown
 <308> L38951
 <400> 448
 CCAGAAAATT GTGCCRAAGA GTTTAGAAAA A

40
 <210> 449
 <211> 31
 <212> DNA
 <213> unknown
 <308> L38951
 <400> 449
 AAGAGTTTAG AAAAYAAAT ATACAATAAA A

45
 <210> 448
 <211> 31
 <212> DNA
 <213> unknown
 <308> L38951
 <400> 448
 CCAGAAAATT GTGCCRAAGA GTTTAGAAAA A

50
 <210> 449
 <211> 31
 <212> DNA
 <213> unknown
 <308> L38951
 <400> 449
 AAGAGTTTAG AAAAYAAAT ATACAATAAA A

55

5
<210> 450
<211> 31
<212> DNA
<213> unknown
<308> M22490
<400> 450
10 CCAACACCGT GAGGASCTTC CACCACGAAG A

15
<210> 451
<211> 31
<212> DNA
<213> unknown
<308> M22632
<400> 451
20 GGGAGTTGGT GCCTAYCGGG ATGATAATGG A

25
<210> 452
<211> 31
<212> DNA
<213> unknown
<308> M22632
<400> 452
30 CTACCGGGAT GATAAYGGAA AGCCTTACGT T

35
<210> 453
<211> 31
<212> DNA
<213> unknown
<308> M22632
<400> 453
40 TGGAAAGCCT TACGTKCTGC CTAGCGTCCG C

45
<210> 454
<211> 31
<212> DNA
<213> unknown
<308> M22632
<400> 454
50 CGGGAGTGGA CCCGCRTCCG GAACAGTGGA A

55
<210> 455
<211> 31
<212> DNA
<213> unknown
<308> M22632
<400> 455
GGGCATTAAT GTTTGYCTCT GCCAATCATA T

5
10
15
20
25
30
35
40
45
50
55

<210> 456
<211> 31
<212> DNA
<213> unknown
<308> M22632
<400> 456
TGCAAGAAGT GAAAGKCATG GCTGACCGCA T

<210> 457
<211> 31
<212> DNA
<213> unknown
<308> M22632
<400> 457
TCAGCCACAG TGTGTRACAC TCAGCATTTG A

<210> 458
<211> 31
<212> DNA
<213> unknown
<308> M22632
<400> 458
CCTTATCAAC AGGCCRCAC T GCAGAAATGA T

<210> 459
<211> 31
<212> DNA
<213> unknown
<308> M22632
<400> 459
TTTTTCTGCT GATGCGYTAC CCTCACCCCTT T

<210> 460
<211> 31
<212> DNA
<213> unknown
<308> M22632
<400> 460
ATGTGGTGTG CAGCCWCTCA TCTCACACTG T

<210> 461
<211> 31
<212> DNA
<213> unknown
<308> M22632
<400> 461
CGGCGCAGGG TGGTASCAGG AAAGAGGATA T

EP 1 024 200 A2

5
 <210> 462
 <211> 31
 <212> DNA
 <213> unknown
 <308> M57567
 <400> 462
 GCTGGGGCTA CAGCAYTTAC GCAGCCGCAC G

10
 <210> 463
 <211> 31
 <212> DNA
 <213> unknown
 <308> M57567
 <400> 463
 TGATGCCCCG AAGCTSCTGC GTGCATCCCC G

15
 <210> 464
 <211> 31
 <212> DNA
 <213> unknown
 <308> M63483
 <400> 464
 GAATCTGAAG ATGAGYTACT TGTAGATGAA G

20
 <210> 465
 <211> 31
 <212> DNA
 <213> unknown
 <308> M63483
 <400> 465
 CCAACAAAGA TACAASTGAA AACGCAGATG G

25
 <210> 466
 <211> 31
 <212> DNA
 <213> unknown
 <308> M63483
 <400> 466
 AGATACAAGT GAAAAYGCAG ATGGTCAAAG T

30
 <210> 467
 <211> 31
 <212> DNA
 <213> unknown
 <308> M87503
 <400> 467
 TGATACAGCT AAGACYATGT TCCGGATTCC C

35
 <210> 468
 <211> 31
 <212> DNA
 <213> unknown
 <308> M87503
 <400> 468
 TGATACAGCT AAGACYATGT TCCGGATTCC C

40
 <210> 469
 <211> 31
 <212> DNA
 <213> unknown
 <308> M87503
 <400> 469
 TGATACAGCT AAGACYATGT TCCGGATTCC C

45
 <210> 470
 <211> 31
 <212> DNA
 <213> unknown
 <308> M87503
 <400> 470
 TGATACAGCT AAGACYATGT TCCGGATTCC C

50
 <210> 471
 <211> 31
 <212> DNA
 <213> unknown
 <308> M87503
 <400> 471
 TGATACAGCT AAGACYATGT TCCGGATTCC C

55
 <210> 472
 <211> 31
 <212> DNA
 <213> unknown
 <308> M87503
 <400> 472
 TGATACAGCT AAGACYATGT TCCGGATTCC C

EP 1 024 200 A2

5
5
10
<210> 468
<211> 31
<212> DNA
<213> unknown
<308> M95627
<400> 468
AGCGAGGTCA CCTTTRCATT GCACTCAGCA T

15
20
<210> 469
<211> 31
<212> DNA
<213> unknown
<308> M95627
<400> 469
CTGCTTACTG ACTACYGGGG CCACACGGCT G

25
30
<210> 470
<211> 31
<212> DNA
<213> unknown
<308> M95627
<400> 470
GCCCTCCAC CCTTGYCCAG ACCTGGTGGG C

35
40
<210> 471
<211> 31
<212> DNA
<213> unknown
<308> U12595
<400> 471
CCAGGCGAAA CCCAGYCTGG AGCTTGCAGG C

45
50
<210> 472
<211> 31
<212> DNA
<213> unknown
<308> U12595
<400> 472
CTGCCAGAAA TGGAGRTTCA CTTGCAGACC A

55
50
55
<210> 473
<211> 31
<212> DNA
<213> unknown
<308> U12595
<400> 473
AGAACCGGGA CAAAARTCAT CATCCACCTG A

EP 1 024 200 A2

```

5      <210> 474
      <211> 31
      <212> DNA
      <213> unknown
      <308> U12595
      <400> 474
      GACCCCAAGG ATGTCSGTGA GTGGCAACAT G

10     <210> 475
      <211> 31
      <212> DNA
      <213> unknown
      <308> U12595
      <400> 475
      CCGCAGCATC TTCTAYGTGC CCGACATGAA A

15     <210> 476
      <211> 31
      <212> DNA
      <213> unknown
      <308> U12595
      <400> 476
      GGTGGACAGT GAGGASATTC CCCTGAACCT C

20     <210> 477
      <211> 31
      <212> DNA
      <213> unknown
      <308> U12595
      <400> 477
      TGCTGGACTT GTTGAYGACC CTAGGGCCAT G

25     <210> 478
      <211> 31
      <212> DNA
      <213> unknown
      <308> U12595
      <400> 478
      GGGCCATGGT GGGCCRCTTG AATGAGCTGC T

30     <210> 479
      <211> 31
      <212> DNA
      <213> unknown
      <308> U15085
      <400> 479
      CTAGGGGCTG GGTATYCCCA CATCACTCAT T

```

EP 1 024 200 A2

5
 <210> 480
 <211> 31
 <212> DNA
 <213> unknown
 <308> U15085
 <400> 480
 GCCCATGTGG AAAGCRCCTG TCTGTTGGAT G

10
 <210> 481
 <211> 31
 <212> DNA
 <213> unknown
 <308> U15085
 <400> 481
 TGGAAAGCAC CTGTCKGTTG GATGATGCTG G

15
 <210> 482
 <211> 31
 <212> DNA
 <213> unknown
 <308> U15085
 <400> 482
 TCTCCTTCAA CAAGGWTCTG CTGACCTGCT G

20
 <210> 483
 <211> 31
 <212> DNA
 <213> unknown
 <308> U15085
 <400> 483
 TTGGGGTGCT GAATARCTTG GCGAATGTCC T

25
 <210> 484
 <211> 31
 <212> DNA
 <213> unknown
 <308> U15085
 <400> 484
 TGCCTCACAG CAGTGMGCAC AAGACTGCCC A

30
 <210> 485
 <211> 31
 <212> DNA
 <213> unknown
 <308> U15085
 <400> 485
 GTGTGGTAGA GCACAYTGGG GCTCCTGAGC C

35
 <210> 486
 <211> 31
 <212> DNA
 <213> unknown
 <308> U15085
 <400> 486
 TGTGCTGCTG GCTGCTGCTG GCTGCTGCTG G

40
 <210> 487
 <211> 31
 <212> DNA
 <213> unknown
 <308> U15085
 <400> 487
 TGTGCTGCTG GCTGCTGCTG GCTGCTGCTG G

45
 <210> 488
 <211> 31
 <212> DNA
 <213> unknown
 <308> U15085
 <400> 488
 TGTGCTGCTG GCTGCTGCTG GCTGCTGCTG G

50
 <210> 489
 <211> 31
 <212> DNA
 <213> unknown
 <308> U15085
 <400> 489
 TGTGCTGCTG GCTGCTGCTG GCTGCTGCTG G

55
 <210> 490
 <211> 31
 <212> DNA
 <213> unknown
 <308> U15085
 <400> 490
 TGTGCTGCTG GCTGCTGCTG GCTGCTGCTG G

5
 <210> 486
 <211> 31
 <212> DNA
 <213> unknown
 <308> U15085
 <400> 486
10 TCTTCTCTCT TGGTGYGATC AGCTGGCGGA G

15
 <210> 487
 <211> 31
 <212> DNA
 <213> unknown
 <308> U15085
 <400> 487
20 ATTCAAACCTC AATGAYGCTA CCATGCCTCT C

25
 <210> 488
 <211> 31
 <212> DNA
 <213> unknown
 <308> U15085
 <400> 488
30 TGTAAAACTT AGCAAYTTGG GGGACCTCAT T

35
 <210> 489
 <211> 31
 <212> DNA
 <213> unknown
 <308> U15085
 <400> 489
40 TCAGTGTGCC ATAGARGACA GCAACTGGTG A

45
 <210> 490
 <211> 31
 <212> DNA
 <213> unknown
 <308> U28386
 <400> 490
50 ACAGAAATGA GCGTYGCAG AATAGAGGTC A

55
 <210> 491
 <211> 31
 <212> DNA
 <213> unknown
 <308> U28386
 <400> 491
60 ATGGAGGTGC CATCCSAGCA TTCATTCTC T

5
5
10
15
20
25
30
35
40
45
50
55

<210> 492
<211> 31
<212> DNA
<213> unknown
<308> U28386
<400> 492
CAGTTCCTGA TATGTMATCT TTAGCATGTG G

<210> 493
<211> 31
<212> DNA
<213> unknown
<308> U28386
<400> 493
TGATGATCCA GAAGTRTTAG CAGATACCTG C

<210> 494
<211> 31
<212> DNA
<213> unknown
<308> U28386
<400> 494
CCCATTCCTT GTCAGYGTTT TCTCTAAGGC A

<210> 495
<211> 31
<212> DNA
<213> unknown
<308> U28386
<400> 495
GGCTGCTGAG AAACTRGGTG AAAC TGAGAA A

<210> 496
<211> 31
<212> DNA
<213> unknown
<308> U28386
<400> 496
GAGGCTTAGA CAAAATGAA GCTCTACAAA A

<210> 497
<211> 31
<212> DNA
<213> unknown
<308> U30255
<400> 497
AACATGAATG ACCACSGCTT TGTGGTCTGT G

5
 <210> 498
 <211> 31
 <212> DNA
 <213> unknown
 <308> U30255
 <400> 498
 TGTCTCCAAA GTTGAYGATT TCTTGGCCAA T

10
 <210> 499
 <211> 31
 <212> DNA
 <213> unknown
 <308> U30255
 <400> 499
 TCTCAAGTTC CAAGAYACCG ATGGCAAACA C

15
 <210> 500
 <211> 31
 <212> DNA
 <213> unknown
 <308> U30255
 <400> 500
 AAGTTCCAAG ACACCRATGG CAAACACCTG C

20
 <210> 501
 <211> 31
 <212> DNA
 <213> unknown
 <308> U30255
 <400> 501
 CACCGTGTCA TCCTCRTCAT ACAATGCCTG A

25
 <210> 502
 <211> 31
 <212> DNA
 <213> unknown
 <308> U45328
 <400> 502
 GGAGTGCGCC ATTCCRGGA AGAAAGGGAC T

30
 <210> 503
 <211> 31
 <212> DNA
 <213> unknown
 <308> U45328
 <400> 503
 CTTGTGGCAT CGTCARAAGG AAGGGATTGG T

35
 <210> 504
 <211> 31
 <212> DNA
 <213> unknown
 <308> U45328
 <400> 504
 TTTGTTGCTT CTTGTTGCTT CTTGTTGCTT C

40
 <210> 505
 <211> 31
 <212> DNA
 <213> unknown
 <308> U45328
 <400> 505
 TTTGTTGCTT CTTGTTGCTT CTTGTTGCTT C

45
 <210> 506
 <211> 31
 <212> DNA
 <213> unknown
 <308> U45328
 <400> 506
 TTTGTTGCTT CTTGTTGCTT CTTGTTGCTT C

50
 <210> 507
 <211> 31
 <212> DNA
 <213> unknown
 <308> U45328
 <400> 507
 TTTGTTGCTT CTTGTTGCTT CTTGTTGCTT C

55
 <210> 508
 <211> 31
 <212> DNA
 <213> unknown
 <308> U45328
 <400> 508
 TTTGTTGCTT CTTGTTGCTT CTTGTTGCTT C

EP 1 024 200 A2

5
 <210> 504
 <211> 31
 <212> DNA
 <213> unknown
 <308> U45328
 <400> 504
 GCGGGTGTGC GGTCTSGATT CGCTGAATTG C

10
 <210> 505
 <211> 31
 <212> DNA
 <213> unknown
 <308> U45328
 <400> 505
 TGTGCGGTCT CGATTSGCTG AATTGCCCGT T

15
 <210> 506
 <211> 31
 <212> DNA
 <213> unknown
 <308> U51127
 <400> 506
 GGATGAAGCC GATCCRGCCA AGTGAAGGC C

20
 <210> 507
 <211> 31
 <212> DNA
 <213> unknown
 <308> U51127
 <400> 507
 ACAGCCCCGA TGAGCWCTG GCTGGCTGCA G

25
 <210> 508
 <211> 31
 <212> DNA
 <213> unknown
 <308> U51127
 <400> 508
 TTCCTGGAAG TGGATYTGGG CCAAGAAGGA G

30
 <210> 509
 <211> 31
 <212> DNA
 <213> unknown
 <308> U88964
 <400> 509
 GAGGGTCCCC AAGGARCATG GCTGGGAGCC G

35
 <210> 508
 <211> 31
 <212> DNA
 <213> unknown
 <308> U51127
 <400> 508
 TTCCTGGAAG TGGATYTGGG CCAAGAAGGA G

40
 <210> 509
 <211> 31
 <212> DNA
 <213> unknown
 <308> U88964
 <400> 509
 GAGGGTCCCC AAGGARCATG GCTGGGAGCC G

45
 <210> 509
 <211> 31
 <212> DNA
 <213> unknown
 <308> U88964
 <400> 509
 GAGGGTCCCC AAGGARCATG GCTGGGAGCC G

50
 <210> 509
 <211> 31
 <212> DNA
 <213> unknown
 <308> U88964
 <400> 509
 GAGGGTCCCC AAGGARCATG GCTGGGAGCC G

55

EP 1 024 200 A2

5
 <210> 510
 <211> 31
 <212> DNA
 <213> unknown
 <308> U88964
 <400> 510
 GAGTGAGCGC CTCCTRCACA AGAGCATCCA G

10
 <210> 511
 <211> 31
 <212> DNA
 <213> unknown
 <308> X15729
 <400> 511
 TTTCCCTGCA AATGTSATGG ATGTTATTGC A

15
 <210> 512
 <211> 31
 <212> DNA
 <213> unknown
 <308> X15729
 <400> 512
 GACAGAATTT CACTGMACCC ACTGCTATTC A

20
 <210> 513
 <211> 31
 <212> DNA
 <213> unknown
 <308> X15729
 <400> 513
 CTCAGGGATG GCCAGKTGCT CTAAGTGGAT T

25
 <210> 514
 <211> 31
 <212> DNA
 <213> unknown
 <308> X15729
 <400> 514
 ATTTGCTTCC TGCCAYTGTC CACATCAATC A

30
 <210> 515
 <211> 31
 <212> DNA
 <213> unknown
 <308> X15729
 <400> 515
 ATTGTGGATG TGTGTSATGA CGTAGAAAAG G

35
 <210> 516
 <211> 31
 <212> DNA
 <213> unknown
 <308> X15729
 <400> 516
 ATTGTGGATG TGTGTSATGA CGTAGAAAAG G

40
 <210> 517
 <211> 31
 <212> DNA
 <213> unknown
 <308> X15729
 <400> 517
 ATTGTGGATG TGTGTSATGA CGTAGAAAAG G

45
 <210> 518
 <211> 31
 <212> DNA
 <213> unknown
 <308> X15729
 <400> 518
 ATTGTGGATG TGTGTSATGA CGTAGAAAAG G

50
 <210> 519
 <211> 31
 <212> DNA
 <213> unknown
 <308> X15729
 <400> 519
 ATTGTGGATG TGTGTSATGA CGTAGAAAAG G

55
 <210> 520
 <211> 31
 <212> DNA
 <213> unknown
 <308> X15729
 <400> 520
 ATTGTGGATG TGTGTSATGA CGTAGAAAAG G

5
10
15
20
25
30
35
40
45
50
55

<210> 516
<211> 31
<212> DNA
<213> unknown
<308> X15729
<400> 516
GTCGAAGACA GAGGTCAGG TCGTTCCAGG G

<210> 517
<211> 31
<212> DNA
<213> unknown
<308> X56253
<400> 517
GCAGTGAGAG AATCCWGGCA GACAGAAGAA A

<210> 518
<211> 31
<212> DNA
<213> unknown
<308> X75861
<400> 518
TGCTTTCATG GGCACRGCAA TGATCTTTAC C

<210> 519
<211> 31
<212> DNA
<213> unknown
<308> X75861
<400> 519
CCTTGTTGAT ACTCARCTCA TTATTGAAAA G

<210> 520
<211> 31
<212> DNA
<213> unknown
<308> X75861
<400> 520
CTCCCTTTT TGTCARCCCC ATCTGTAGCC T

<210> 521
<211> 31
<212> DNA
<213> unknown
<308> X75861
<400> 521
CCCTTCCTTC CTCATYGTTG TTTGGTATGC G

5
10
15
20
25
30
35
40
45
50
55

<210> 522
<211> 31
<212> DNA
<213> unknown
<308> X75861
<400> 522
CCATTGGAAA GAAGTKCAGT GGTCCCATTG T

<210> 523
<211> 31
<212> DNA
<213> unknown
<308> D13900
<400> 523
GAGAGCCATG GCCGCYCTGC GTGTCCTGCT G

<210> 524
<211> 31
<212> DNA
<213> unknown
<308> D13900
<400> 524
TCCTGCTGTC CTGCGYCCGC GGCCCGCTGA G

<210> 525
<211> 31
<212> DNA
<213> unknown
<308> D13900
<400> 525
ACCAGGCCCT GAAGAYCTTC GAGGAGGACC C

<210> 526
<211> 31
<212> DNA
<213> unknown
<308> D13900
<400> 526
CTTCGAGGAG GACCCRGCCG TTGGGGGCAT T

<210> 527
<211> 31
<212> DNA
<213> unknown
<308> D13900
<400> 527
GGAGATGGTC CTCACYGGTG ACGCGATCTC A

5
 <210> 528
 <211> 31
 <212> DNA
 <213> unknown
 <308> D13900
 <400> 528
 CCGTGGTGTG CAGTTYCTCT CCAATTGCTG C

10
 <210> 529
 <211> 31
 <212> DNA
 <213> unknown
 <308> D13900
 <400> 529
 GGAAAGCGGG GCCTGYGGGT CCTTGTGTCC C

15
 <210> 530
 <211> 31
 <212> DNA
 <213> unknown
 <308> D14812
 <400> 530
 TGTTTCAGTT TGTTYCCGT TCTGTTTAA A

20
 <210> 531
 <211> 31
 <212> DNA
 <213> unknown
 <308> D14812
 <400> 531
 GTTCCTATCC CTTTGMCCT ACACAATTTT C

25
 <210> 532
 <211> 31
 <212> DNA
 <213> unknown
 <308> D14812
 <400> 532
 AGGCATTCTA TCTTTSCTCA AATTGTTGAA G

30
 <210> 533
 <211> 31
 <212> DNA
 <213> unknown
 <308> D28473
 <400> 533
 GCTGGCCCGC GGGGCWGTCC AGGCACGCGA G

35
 <210> 534
 <211> 31
 <212> DNA
 <213> unknown
 <308> D28473
 <400> 534
 GCTGGCCCGC GGGGCWGTCC AGGCACGCGA G

40
 <210> 535
 <211> 31
 <212> DNA
 <213> unknown
 <308> D28473
 <400> 535
 GCTGGCCCGC GGGGCWGTCC AGGCACGCGA G

45
 <210> 536
 <211> 31
 <212> DNA
 <213> unknown
 <308> D28473
 <400> 536
 GCTGGCCCGC GGGGCWGTCC AGGCACGCGA G

50
 <210> 537
 <211> 31
 <212> DNA
 <213> unknown
 <308> D28473
 <400> 537
 GCTGGCCCGC GGGGCWGTCC AGGCACGCGA G

55
 <210> 538
 <211> 31
 <212> DNA
 <213> unknown
 <308> D28473
 <400> 538
 GCTGGCCCGC GGGGCWGTCC AGGCACGCGA G

5
10
15
20
25
30
35
40
45
50
55

<210> 534
<211> 31
<212> DNA
<213> unknown
<308> D28473
<400> 534
AAGATTGGA TGGGAKTGCC ATGGCTTACC T

<210> 535
<211> 31
<212> DNA
<213> unknown
<308> D28473
<400> 535
GGACGATTAC TCATTKTAAT GGAAGCCAGA T

<210> 536
<211> 31
<212> DNA
<213> unknown
<308> D28473
<400> 536
AGAACTGTCA GGAGCRAAGA TCTCAGATCT C

<210> 537
<211> 31
<212> DNA
<213> unknown
<308> D28473
<400> 537
TCTACAATGA GAACAYGGTT AGAGAAAGCC C

<210> 538
<211> 31
<212> DNA
<213> unknown
<308> D28473
<400> 538
AGATAAAAAC AAGTAYGGCA TTCGGCTAAG G

<210> 539
<211> 31
<212> DNA
<213> unknown
<308> D28473
<400> 539
AGATGAAGGA ATGGCKCGGG AAGTCATCAA T

EP 1 024 200 A2

5
 <210> 540
 <211> 31
 <212> DNA
 <213> unknown
 <308> D28473
 <400> 540
 AAGCAAAGTC TGAAGKAACA TATCTGAATA G

10
 <210> 541
 <211> 31
 <212> DNA
 <213> unknown
 <308> D28473
 <400> 541
 TTGAAACCAT ATCCARTTTC TCCATCGGAT A

15
 <210> 542
 <211> 31
 <212> DNA
 <213> unknown
 <308> D28473
 <400> 542
 CAGCTCCTGA ATGCARAGCC ACAAGAGTGT T

20
 <210> 543
 <211> 31
 <212> DNA
 <213> unknown
 <308> D28473
 <400> 543
 GGTGTGTGTA AAATAYGCTG CTTGGATTGA A

25
 <210> 544
 <211> 31
 <212> DNA
 <213> unknown
 <308> D28473
 <400> 544
 GTCAAGTAAC TTGAGWCCTC ACAGTAATCT T

30
 <210> 545
 <211> 31
 <212> DNA
 <213> unknown
 <308> D32050
 <400> 545
 TTCAGAAGGG CACAGRTGCC CGACCATACA C

35
 <210> 546
 <211> 31
 <212> DNA
 <213> unknown
 <308> D32050
 <400> 546
 TTTGAGTGGT TTTGAGTGGT TTTGAGTGGT T

40
 <210> 547
 <211> 31
 <212> DNA
 <213> unknown
 <308> D32050
 <400> 547
 TTTGAGTGGT TTTGAGTGGT TTTGAGTGGT T

45
 <210> 548
 <211> 31
 <212> DNA
 <213> unknown
 <308> D32050
 <400> 548
 TTTGAGTGGT TTTGAGTGGT TTTGAGTGGT T

50
 <210> 549
 <211> 31
 <212> DNA
 <213> unknown
 <308> D32050
 <400> 549
 TTTGAGTGGT TTTGAGTGGT TTTGAGTGGT T

55
 <210> 550
 <211> 31
 <212> DNA
 <213> unknown
 <308> D32050
 <400> 550
 TTTGAGTGGT TTTGAGTGGT TTTGAGTGGT T

5
10
15
20
25
30
35
40
45
50
55

<210> 546
<211> 31
<212> DNA
<213> unknown
<308> D32050
<400> 546
GGTGCTGGCT GACCAYGCTC GGACCATCAC T

<210> 547
<211> 31
<212> DNA
<213> unknown
<308> D32050
<400> 547
TCCCGGAGAC ACTGCTTGGC TCCTCTATGA C

<210> 548
<211> 31
<212> DNA
<213> unknown
<308> D32050
<400> 548
CCAAACAAGG ATGTGSAGAG GGAGATCGCT G

<210> 549
<211> 31
<212> DNA
<213> unknown
<308> D32050
<400> 549
GTGCCTGTGT CAAGTYCCCC AGAATGCAGC C

<210> 550
<211> 31
<212> DNA
<213> unknown
<308> D32050
<400> 550
TCGCATCTAT AGATAMCGGC TCTCCAGACC T

<210> 551
<211> 31
<212> DNA
<213> unknown
<308> D78151
<400> 551
GGGAGGCCGG GACAARGCGC CGGTGCAGCC C

EP 1 024 200 A2

5
 <210> 552
 <211> 31
 <212> DNA
 <213> unknown
 <308> D78151
 <400> 552
 CCACAATGCA GAGCAKGAGG CTTGCGACCT G

10
 <210> 553
 <211> 31
 <212> DNA
 <213> unknown
 <308> D78151
 <400> 553
 TGAAGATGTC GAGGAKTATG AGGACCTGAC A

20
 <210> 554
 <211> 31
 <212> DNA
 <213> unknown
 <308> D78151
 <400> 554
 CCTCCTCTTT TGTGARTGGC TTTGTGAATG C

30
 <210> 555
 <211> 31
 <212> DNA
 <213> unknown
 <308> D78151
 <400> 555
 TCCGGAATGA GTGTGWCCCT GCTCTGGCAC T

35
 <210> 556
 <211> 31
 <212> DNA
 <213> unknown
 <308> D78151
 <400> 556
 TGCTCTCAGA CTATGWTCTC CACAACAGCA A

45
 <210> 557
 <211> 31
 <212> DNA
 <213> unknown
 <308> D78151
 <400> 557
 CAGTTTTGCC AACACRCTGG TGGATGTGTG T

55

5
 <210> 558
 <211> 31
 <212> DNA
 <213> unknown
 <308> D78151
 <400> 558
 AGATATGGGG AGCCTRCACCT CCGGAGGGCT G

10
 <210> 559
 <211> 31
 <212> DNA
 <213> unknown
 <308> D78151
 <400> 559
 CGTCTGGCTG CAATGYTGCG CCAGTTAGCT C

15
 <210> 560
 <211> 31
 <212> DNA
 <213> unknown
 <308> D78151
 <400> 560
 ATACAACCCC AGTGTGTTG GCCCACGGGG A

20
 <210> 561
 <211> 31
 <212> DNA
 <213> unknown
 <308> D78151
 <400> 561
 AGGGGCTCTG AACTGYAGCT GATGTTATCA G

25
 <210> 562
 <211> 31
 <212> DNA
 <213> unknown
 <308> D78151
 <400> 562
 GATGTTATCA GCAGGMCATG CATCCTGCTG C

30
 <210> 563
 <211> 31
 <212> DNA
 <213> unknown
 <308> J04611
 <400> 563
 TGGTTGATGC CTCCARGGCT ATGTTTGAAT C

35
 <210> 564
 <211> 31
 <212> DNA
 <213> unknown
 <308> J04611
 <400> 564
 TGGTTGATGC CTCCARGGCT ATGTTTGAAT C

40
 <210> 565
 <211> 31
 <212> DNA
 <213> unknown
 <308> J04611
 <400> 565
 TGGTTGATGC CTCCARGGCT ATGTTTGAAT C

45
 <210> 566
 <211> 31
 <212> DNA
 <213> unknown
 <308> J04611
 <400> 566
 TGGTTGATGC CTCCARGGCT ATGTTTGAAT C

50
 <210> 567
 <211> 31
 <212> DNA
 <213> unknown
 <308> J04611
 <400> 567
 TGGTTGATGC CTCCARGGCT ATGTTTGAAT C

55
 <210> 568
 <211> 31
 <212> DNA
 <213> unknown
 <308> J04611
 <400> 568
 TGGTTGATGC CTCCARGGCT ATGTTTGAAT C

EP 1 024 200 A2

5
 <210> 564
 <211> 31
 <212> DNA
 <213> unknown
 <308> J04611
 <400> 564
 GTGATCTCTG TGGGCRTTTA TAATCTGGTC C

10
 <210> 565
 <211> 31
 <212> DNA
 <213> unknown
 <308> J04611
 <400> 565
 CCTAGCGATA CCAAGMGGTC TCAGATCTAT G

15
 <210> 566
 <211> 31
 <212> DNA
 <213> unknown
 <308> J04611
 <400> 566
 TACTGCTGAA GAAACRCCAT TACCTGAGGC C

20
 <210> 567
 <211> 31
 <212> DNA
 <213> unknown
 <308> J04611
 <400> 567
 AAAAGACTGG GCTCCYTGGT GGATGAGTTT A

25
 <210> 568
 <211> 31
 <212> DNA
 <213> unknown
 <308> J04611
 <400> 568
 TTTAAGGAGC TTGTTYACCC ACCAGATTAC A

30
 <210> 569
 <211> 31
 <212> DNA
 <213> unknown
 <308> J04611
 <400> 569
 CGGGCTGAAG AGTGGKCTGA AGAAGCAGGA G

35
 <210> 569
 <211> 31
 <212> DNA
 <213> unknown
 <308> J04611
 <400> 569
 CGGGCTGAAG AGTGGKCTGA AGAAGCAGGA G

40
 <210> 569
 <211> 31
 <212> DNA
 <213> unknown
 <308> J04611
 <400> 569
 CGGGCTGAAG AGTGGKCTGA AGAAGCAGGA G

45
 <210> 569
 <211> 31
 <212> DNA
 <213> unknown
 <308> J04611
 <400> 569
 CGGGCTGAAG AGTGGKCTGA AGAAGCAGGA G

50
 <210> 569
 <211> 31
 <212> DNA
 <213> unknown
 <308> J04611
 <400> 569
 CGGGCTGAAG AGTGGKCTGA AGAAGCAGGA G

55

5
 <210> 570
 <211> 31
 <212> DNA
 <213> unknown
 <308> L15189
 <400> 570
 GAGGCTCATG GGAAAYTGTA TTCTCCGAGT C

10
 <210> 571
 <211> 31
 <212> DNA
 <213> unknown
 <308> L15189
 <400> 571
 TTGGAGCATT TGTGTYGATG AAGATGAAAG A

20
 <210> 572
 <211> 31
 <212> DNA
 <213> unknown
 <308> L15189
 <400> 572
 AGGGGTTGAT TTGACYAAAG ACAACATGGC A

30
 <210> 573
 <211> 31
 <212> DNA
 <213> unknown
 <308> L15189
 <400> 573
 TGCTGAAAAG GCTAARTGTG AACTCTCCTC A

35
 <210> 574
 <211> 31
 <212> DNA
 <213> unknown
 <308> L15189
 <400> 574
 CCAGTCTTCT GGTGGRTTAA GCAAAGATGA T

45
 <210> 575
 <211> 31
 <212> DNA
 <213> unknown
 <308> L15189
 <400> 575
 GAAGAGATTT CCAAARTGAG GGAGCTCCTG G

50

55

EP 1 024 200 A2

5
 <210> 576
 <211> 31
 <212> DNA
 <213> unknown
 <308> L15189
 <400> 576
 CTTCAGCAGG CATCAYTGAA GCTGTTCGAA A

10
 <210> 577
 <211> 31
 <212> DNA
 <213> unknown
 <308> M29877
 <400> 577
 GGATGAGGTC GCGGCSGGCG GGTCCCGCGC T

15
 <210> 578
 <211> 31
 <212> DNA
 <213> unknown
 <308> M29877
 <400> 578
 CGACTTCGGA CCGCARTTCA CTGCGCGCTT C

20
 <210> 579
 <211> 31
 <212> DNA
 <213> unknown
 <308> M29877
 <400> 579
 CAAGTATGTA GTTTTSACGA CAAAGCATCA C

25
 <210> 580
 <211> 31
 <212> DNA
 <213> unknown
 <308> M29877
 <400> 580
 GCTTCACAAA CTGGCYGAGT CCTGTGTCTT G

30
 <210> 581
 <211> 31
 <212> DNA
 <213> unknown
 <308> M29877
 <400> 581
 ATAAATTCAA GCCACRGAGC TTGCCAGATC A

35
 <210> 581
 <211> 31
 <212> DNA
 <213> unknown
 <308> M29877
 <400> 581
 ATAAATTCAA GCCACRGAGC TTGCCAGATC A

40
 <210> 581
 <211> 31
 <212> DNA
 <213> unknown
 <308> M29877
 <400> 581
 ATAAATTCAA GCCACRGAGC TTGCCAGATC A

45
 <210> 581
 <211> 31
 <212> DNA
 <213> unknown
 <308> M29877
 <400> 581
 ATAAATTCAA GCCACRGAGC TTGCCAGATC A

50
 <210> 581
 <211> 31
 <212> DNA
 <213> unknown
 <308> M29877
 <400> 581
 ATAAATTCAA GCCACRGAGC TTGCCAGATC A

55

EP 1 024 200 A2

5
 <210> 582
 <211> 31
 <212> DNA
 <213> unknown
 <308> M97935
 <400> 582
 AAGCTTCTTG GTCCTMACGC CAGCCCCGAT G

10
 <210> 583
 <211> 31
 <212> DNA
 <213> unknown
 <308> M97935
 <400> 583
 AGTCTATTAG CCACARAATT GGGAAAGGAG T

15
 <210> 584
 <211> 31
 <212> DNA
 <213> unknown
 <308> M97935
 <400> 584
 TGTATTGCAT TAAATRTAAT ATCGACACAG T

20
 <210> 585
 <211> 31
 <212> DNA
 <213> unknown
 <308> U09587
 <400> 585
 CCAGCTTGAT AACTAYGGAC AGCAAGAACT T

25
 <210> 586
 <211> 31
 <212> DNA
 <213> unknown
 <308> U09587
 <400> 586
 AGAAATTGAG CACTTYGTAG ATCCCAGTGA G

30
 <210> 587
 <211> 31
 <212> DNA
 <213> unknown
 <308> U09587
 <400> 587
 GAAATGGAGA TGCTGSTGAA TGAGAAAGGG G

35
 <210> 588
 <211> 31
 <212> DNA
 <213> unknown
 <308> U09587
 <400> 588
 GAAATGGAGA TGCTGSTGAA TGAGAAAGGG G

40
 <210> 589
 <211> 31
 <212> DNA
 <213> unknown
 <308> U09587
 <400> 589
 GAAATGGAGA TGCTGSTGAA TGAGAAAGGG G

45
 <210> 590
 <211> 31
 <212> DNA
 <213> unknown
 <308> U09587
 <400> 590
 GAAATGGAGA TGCTGSTGAA TGAGAAAGGG G

50
 <210> 591
 <211> 31
 <212> DNA
 <213> unknown
 <308> U09587
 <400> 591
 GAAATGGAGA TGCTGSTGAA TGAGAAAGGG G

55
 <210> 592
 <211> 31
 <212> DNA
 <213> unknown
 <308> U09587
 <400> 592
 GAAATGGAGA TGCTGSTGAA TGAGAAAGGG G

5
 <210> 588
 <211> 31
 <212> DNA
 <213> unknown
 <308> U09587
 <400> 588
 CTGCTGAATG AGAAARGGGA ATTCACAATT G

10
 <210> 589
 <211> 31
 <212> DNA
 <213> unknown
 <308> U09587
 <400> 589
 TTTCCCTGCT GTAGTYGCTC CATTCAAATG T

15
 <210> 590
 <211> 31
 <212> DNA
 <213> unknown
 <308> U09587
 <400> 590
 CATTCAAATG TTCCGYCCTC CCACTGAGCC A

20
 <210> 591
 <211> 31
 <212> DNA
 <213> unknown
 <308> U09587
 <400> 591
 CAATGGCAAC ATCACRTGGG CTGATGTGGA G

25
 <210> 592
 <211> 31
 <212> DNA
 <213> unknown
 <308> U41371
 <400> 592
 GAGCTTCAGG CCAAGKTGGC AGAGATCGGA G

30
 <210> 593
 <211> 31
 <212> DNA
 <213> unknown
 <308> U41371
 <400> 593
 AGGATCTTTG AGGCTKTAA GCTCACTGAT G

35
 <210> 594
 <211> 31
 <212> DNA
 <213> unknown
 <308> U41371
 <400> 594
 AGGATCTTTG AGGCTKTAA GCTCACTGAT G

40
 <210> 595
 <211> 31
 <212> DNA
 <213> unknown
 <308> U41371
 <400> 595
 AGGATCTTTG AGGCTKTAA GCTCACTGAT G

45
 <210> 596
 <211> 31
 <212> DNA
 <213> unknown
 <308> U41371
 <400> 596
 AGGATCTTTG AGGCTKTAA GCTCACTGAT G

50
 <210> 597
 <211> 31
 <212> DNA
 <213> unknown
 <308> U41371
 <400> 597
 AGGATCTTTG AGGCTKTAA GCTCACTGAT G

55
 <210> 598
 <211> 31
 <212> DNA
 <213> unknown
 <308> U41371
 <400> 598
 AGGATCTTTG AGGCTKTAA GCTCACTGAT G

5
10
<210> 594
<211> 31
<212> DNA
<213> unknown
<308> U54778
<400> 594
TGAGACTGAG CTAAARTTAA TCTGTTGTGA C

15
20
<210> 595
<211> 31
<212> DNA
<213> unknown
<308> U54778
<400> 595
TTTTTCAGCA TTACTRTGTC GTCTCCGTGC C

25
30
<210> 596
<211> 31
<212> DNA
<213> unknown
<308> U54778
<400> 596
AAGCCATTAG ACTTAYAGGT GATGCAAGCA T

35
40
<210> 597
<211> 31
<212> DNA
<213> unknown
<308> U54778
<400> 597
ACACTATAGA GGCATRAGTG GTATCAGTTT T

45
50
<210> 598
<211> 31
<212> DNA
<213> unknown
<308> U54778
<400> 598
TAATTGGGTG TTAGCYTGAG ATGGTTAAAG G

55
<210> 599
<211> 31
<212> DNA
<213> unknown
<308> U54778
<400> 599
ACAGATACAG AATCTKTATT GTTCTTACTG A

EP 1 024 200 A2

5
 <210> 600
 <211> 31
 <212> DNA
 <213> unknown
 <308> U57650
 <400> 600
 GAGCCTCTCC GAGACRTTGT TCCAGCGACT G

10
 <210> 601
 <211> 31
 <212> DNA
 <213> unknown
 <308> U57650
 <400> 601
 GACAAGCCTG TTGTCRTCCA TTGAAGACAA G

20
 <210> 602
 <211> 31
 <212> DNA
 <213> unknown
 <308> U57650
 <400> 602
 GAAAATCCTG CAGCTSATTA AGTCACAGAA A

30
 <210> 603
 <211> 31
 <212> DNA
 <213> unknown
 <308> U57650
 <400> 603
 TCCCAGGACA CAGGARTCAA GGCCCAGTGA C

35
 <210> 604
 <211> 31
 <212> DNA
 <213> unknown
 <308> U57650
 <400> 604
 AACACCGAGC TCCCGYATCA CGGCAAGCAC C

40
 <210> 605
 <211> 31
 <212> DNA
 <213> unknown
 <308> U57650
 <400> 605
 CCATGCAGTG AAGCCYTCAG TGAGCTGCCA C

45
 <210> 606
 <211> 31
 <212> DNA
 <213> unknown
 <308> U57650
 <400> 606
 CCATGCAGTG AAGCCYTCAG TGAGCTGCCA C

50
 <210> 607
 <211> 31
 <212> DNA
 <213> unknown
 <308> U57650
 <400> 607
 CCATGCAGTG AAGCCYTCAG TGAGCTGCCA C

55
 <210> 608
 <211> 31
 <212> DNA
 <213> unknown
 <308> U57650
 <400> 608
 CCATGCAGTG AAGCCYTCAG TGAGCTGCCA C

5
<210> 606
<211> 31
<212> DNA
<213> unknown
<308> U57650
<400> 606
TCTGGGTCCC CAGCTYGCTC TTGGTACTTG G

10
<210> 607
<211> 31
<212> DNA
<213> unknown
15 <308> U57650
<400> 607
GGGACCCAG TGCCTYGTTG AGGGCGCCAT T

20
<210> 608
<211> 31
<212> DNA
<213> unknown
<308> U57650
25 <400> 608
GATCGAACAC TCATGRTGTG CCAAGTGCTG T

30
<210> 609
<211> 31
<212> DNA
<213> unknown
<308> U57650
35 <400> 609
GCTTTAGCTA AAGTCYCGCG GGTTCGGCA T

40
<210> 610
<211> 31
<212> DNA
<213> unknown
<308> U80040
45 <400> 610
GAACCGGCCG CTGACMCTCT CGGAGAAGAT T

50
<210> 611
<211> 31
<212> DNA
<213> unknown
<308> U80040
55 <400> 611
GGGATCCCCCT GGGAGYTGAA GTGCCCCAAG G

5
<210> 612
<211> 31
<212> DNA
<213> unknown
<308> U80040
<400> 612
AATCTGCAAC ATGGGYGCAG AAATTGGGGC C

10
<210> 613
<211> 31
<212> DNA
<213> unknown
15
<308> U80040
<400> 613
CCAACAACCT GCTCAYTGGT GCCATCAACA T

20
<210> 614
<211> 31
<212> DNA
<213> unknown
<308> U80040
25
<400> 614
GATTCTAGAG AACCTKTTGT TCTTGCAAGG A

30
<210> 615
<211> 31
<212> DNA
<213> unknown
<308> U83843
35
<400> 615
ATTCTGAAAC TTCTTSATGT TGTCCATCCT G

40
<210> 616
<211> 31
<212> DNA
<213> unknown
<308> U83843
45
<400> 616
TGAGGATTAT CAGGCRATTG TTGATGCTGA G

50
<210> 617
<211> 31
<212> DNA
<213> unknown
<308> U83843
55
<400> 617
GGCAATTGTT GATGCGYAGT GGAACATTCT C

EP 1 024 200 A2

5
 <210> 618
 <211> 31
 <212> DNA
 <213> unknown
 <308> U83843
 <400> 618
 GGATCTGAAG AGGACRATGA TGGCCTGTGG A

10
 <210> 619
 <211> 31
 <212> DNA
 <213> unknown
 <308> U83843
 <400> 619
 GGAGACAGAG CGGTCYCTGC ATGATGCCAT C

15
 <210> 620
 <211> 31
 <212> DNA
 <213> unknown
 <308> U91932
 <400> 620
 AGTGCCCACC CGGTCSGCCC GGCACAGCCA T

20
 <210> 621
 <211> 31
 <212> DNA
 <213> unknown
 <308> X54942
 <400> 621
 TGAGAAATGT ACAAAYCTTT CATCCATACC T

25
 <210> 622
 <211> 31
 <212> DNA
 <213> unknown
 <308> X54942
 <400> 622
 AGAAATGTAC AAATCYTTCA TCCATACCTG T

30
 <210> 623
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59417
 <400> 623
 TTAAAGTAG TGCTTSTACC AACATGTCCC G

35
 <210> 624
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59417
 <400> 624
 TTTAAAGTAG TGCTTSTACC AACATGTCCC G

40
 <210> 625
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59417
 <400> 625
 TTTAAAGTAG TGCTTSTACC AACATGTCCC G

45
 <210> 626
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59417
 <400> 626
 TTTAAAGTAG TGCTTSTACC AACATGTCCC G

50
 <210> 627
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59417
 <400> 627
 TTTAAAGTAG TGCTTSTACC AACATGTCCC G

55
 <210> 628
 <211> 31
 <212> DNA
 <213> unknown
 <308> X59417
 <400> 628
 TTTAAAGTAG TGCTTSTACC AACATGTCCC G

5
5
10
15
20
25
30
35
40
45
50
55

<210> 624
<211> 31
<212> DNA
<213> unknown
<308> X59417
<400> 624
AAGGCCCTCA GGTATRTAAG TGTGATCCTG C

<210> 625
<211> 31
<212> DNA
<213> unknown
<308> X59417
<400> 625
TGATTGGACA TTTGARCAGA CAGTGGAAC T

<210> 626
<211> 31
<212> DNA
<213> unknown
<308> X59417
<400> 626
AGTTGAAAAT CCTAARTTCA GGATTCTTAC A

<210> 627
<211> 31
<212> DNA
<213> unknown
<308> X64364
<400> 627
AGCCTCCGGG GCTGCGGGCA CAGTCTTCAC T

<210> 628
<211> 31
<212> DNA
<213> unknown
<308> X64364
<400> 628
GTTCAAGGTG GACTCSGACG ACCAGTGGGG A

<210> 629
<211> 31
<212> DNA
<213> unknown
<308> X64364
<400> 629
TCCCAGAGTG AAGGCGGTGA AGTCGTCAGA A

5
 <210> 630
 <211> 31
 <212> DNA
 <213> unknown
 <308> X64364
 <400> 630
 10 GACGCTCCCT GCTCCRCGTC TGC GCCGCCG C

15
 <210> 631
 <211> 31
 <212> DNA
 <213> unknown
 <308> X64364
 <400> 631
 ACCGAGGGCG ACCCCRTCAC AGCCTCAAGT C

20
 <210> 632
 <211> 31
 <212> DNA
 <213> unknown
 <308> X64364
 25 <400> 632
 GGACGGCCGG CTCTCYATAG CACCAGGGCT C

30
Claims

- 35 1. A nucleic acid segment of between 10 and 100 contiguous bases selected from a fragment shown in TABLE 1, wherein said segment comprises a polymorphic site or an immediately adjacent base, or the complement of said segment.
- 40 2. The segment of claim 1 that is DNA.
3. The segment of claim 1 that is RNA.
4. The segment of claim 1 that is less than 50 bases.
5. The segment of claim 1 that is less than 20 bases.
- 45 6. The segment of claim 1, wherein the fragment is D42043 and the polymorphic site is at position 1470.
7. The segment of claim 1, wherein the polymorphic site is diallelic.
- 50 8. The segment of claim 1, wherein the polymorphic form occupying the polymorphic site is the reference base for the fragment listed in TABLE 1, column 4.
9. The segment of claim 1, wherein the polymorphic form occupying the polymorphic site is an alternative form for the fragment listed in TABLE 1, column 5.
- 55 10. An allele-specific oligonucleotide that hybridizes to a segment of a fragment shown in TABLE 1, column 6 or its complement.
11. The allele-specific oligonucleotide of claim 10 that is probe.

12. The allele-specific oligonucleotide of claim 10, wherein a central position of the probe aligns with the polymorphic site of the fragment.
13. The allele-specific oligonucleotide of claim 10 that is a primer.
14. The allele-specific oligonucleotide of claim 13, wherein the 3' end of the primer aligns with the polymorphic site or an immediately adjacent base of the fragment.
15. An isolated nucleic acid comprising a sequence of TABLE 1, column 6 or the complement thereof, wherein the polymorphic site within the sequence or complement is occupied by a base other than the reference base shown in TABLE 1, column 4.
16. A method of analyzing a nucleic acid, comprising:
 - obtaining the nucleic acid from an individual; and
 - determining a base occupying any one of the polymorphic sites shown in TABLE 1.
17. The method of claim 16, wherein the determining comprises determining a set of bases occupying a set of the polymorphic sites shown in TABLE 1.
18. The method of claim 16, wherein the nucleic acid is obtained from a plurality of individuals, and a base occupying one of the polymorphic positions is determined in each of the individuals, and the method further comprising testing each individual for the presence of a disease phenotype, and correlating the presence of the disease phenotype with the base.
19. The nucleic acid of claim 15 that is DNA.
20. The nucleic acid of claim 15 that is RNA.
21. The nucleic acid of claim 15 that is less than 50 bases.
22. The nucleic acid of claim 15 that is less than 20 bases.
23. The nucleic acid of claim 15, wherein the fragment is AF005058 and the polymorphic site is at position 3583.
24. The nucleic acid of claim 15, wherein the polymorphic site is diallelic.
25. The nucleic acid of claim 15, wherein the base occupying the polymorphic site is the correlating base for the fragment listed in TABLE 1, column 4.
26. The nucleic acid of claim 15, wherein the base occupying the polymorphic site is the correlating base for the fragment listed in TABLE 1, column 5.
27. The segment of claim 1, wherein said fragment is selected from TABLE 1, SEQ ID Nos: 1-10.
28. The segment of claim 1, wherein said fragment is selected from TABLE 1, SEQ ID Nos: 11-20.
29. The segment of claim 1, wherein said fragment is selected from TABLE 1, SEQ ID Nos:21-30.
30. The segment of claim 1, wherein said fragment is selected from TABLE 1, SEQ ID NOS:31-40.
31. The segment of claim 1, wherein said fragment is selected from TABLE 1, SEQ ID NOS:41-50.
32. The segment of claim 1, wherein said fragment is selected from TABLE 1, SEQ ID NOS:51-60.
33. The segment of claim 1, wherein said fragment is selected from TABLE 1, SEQ ID NOS:61-70.

34. The segment of claim 1, wherein said fragment is selected from TABLE 1, SEQ ID NOS:71-80.

35. The segment of claim 1, wherein said fragment is selected from TABLE 1, SEQ ID NOS:81-90.

5 36. The segment of claim 1, wherein said fragment is selected from TABLE 1, SEQ ID NOS:91-100.

37. The segment of claim 1, wherein said fragment is selected from TABLE 1, SEQ ID NOS:101-110.

38. The segment of claim 1, wherein said fragment is selected from TABLE 1, SEQ ID NOS:111-120.

10 39. The segment of claim 1, wherein said fragment is selected from TABLE 1, SEQ ID NOS:121-130.

40. The segment of claim 1, wherein said fragment is selected from TABLE 1, SEQ ID NOS:131-140.

15 41. The segment of claim 1, wherein said fragment is selected from TABLE 1, SEQ ID NOS:141-150.

42. The segment of claim 1, wherein said fragment is selected from TABLE 1, SEQ ID NOS:151-158.

43. The segment of claim 15, wherein said fragment is selected from TABLE 1, SEQ ID NOS: 1-10.

20 44. The segment of claim 15, wherein said fragment is selected from TABLE 1, SEQ ID NOS: 11-20.

45. The nucleic acid of claim 15, wherein said fragment is selected from TABLE 1, SEQ ID NOS: 21-30.

25 46. The nucleic acid of claim 15, wherein said fragment is selected from TABLE 1, SEQ ID NOS:31-40.

47. The nucleic acid of claim 15, wherein said fragment is selected from TABLE 1, SEQ ID NOS:41-50.

48. The nucleic acid of claim 15, wherein said fragment is selected from TABLE 1, SEQ ID NOS:51-60.

30 49. The nucleic acid of claim 15, wherein said fragment is selected from TABLE 1, SEQ ID NOS:61-70.

50. The nucleic acid of claim 15, wherein said fragment is selected from TABLE 1, SEQ ID NOS:71-80.

35 51. The nucleic acid of claim 15, wherein said fragment is selected from TABLE 1, SEQ ID NOS:81-90.

52. The nucleic acid of claim 15, wherein said fragment is selected from TABLE 1, SEQ ID NOS:91-100.

53. The nucleic acid of claim 15, wherein said fragment is selected from TABLE 1, SEQ ID NOS:101-110.

40 54. The nucleic acid of claim 15, wherein said fragment is selected from TABLE 1, SEQ ID NOS:111-120.

55. The nucleic acid of claim 15, wherein said fragment is selected from TABLE 1, SEQ ID NOS:121-130.

45 56. The nucleic acid of claim 15, wherein said fragment is selected from TABLE 1, SEQ ID NOS:131-140.

57. The nucleic acid of claim 15, wherein said fragment is selected from TABLE 1, SEQ ID NOS:141-150.

50 58. The nucleic acid of claim 15, wherein said fragment is selected from TABLE 1, SEQ ID NOS:151-158.

55

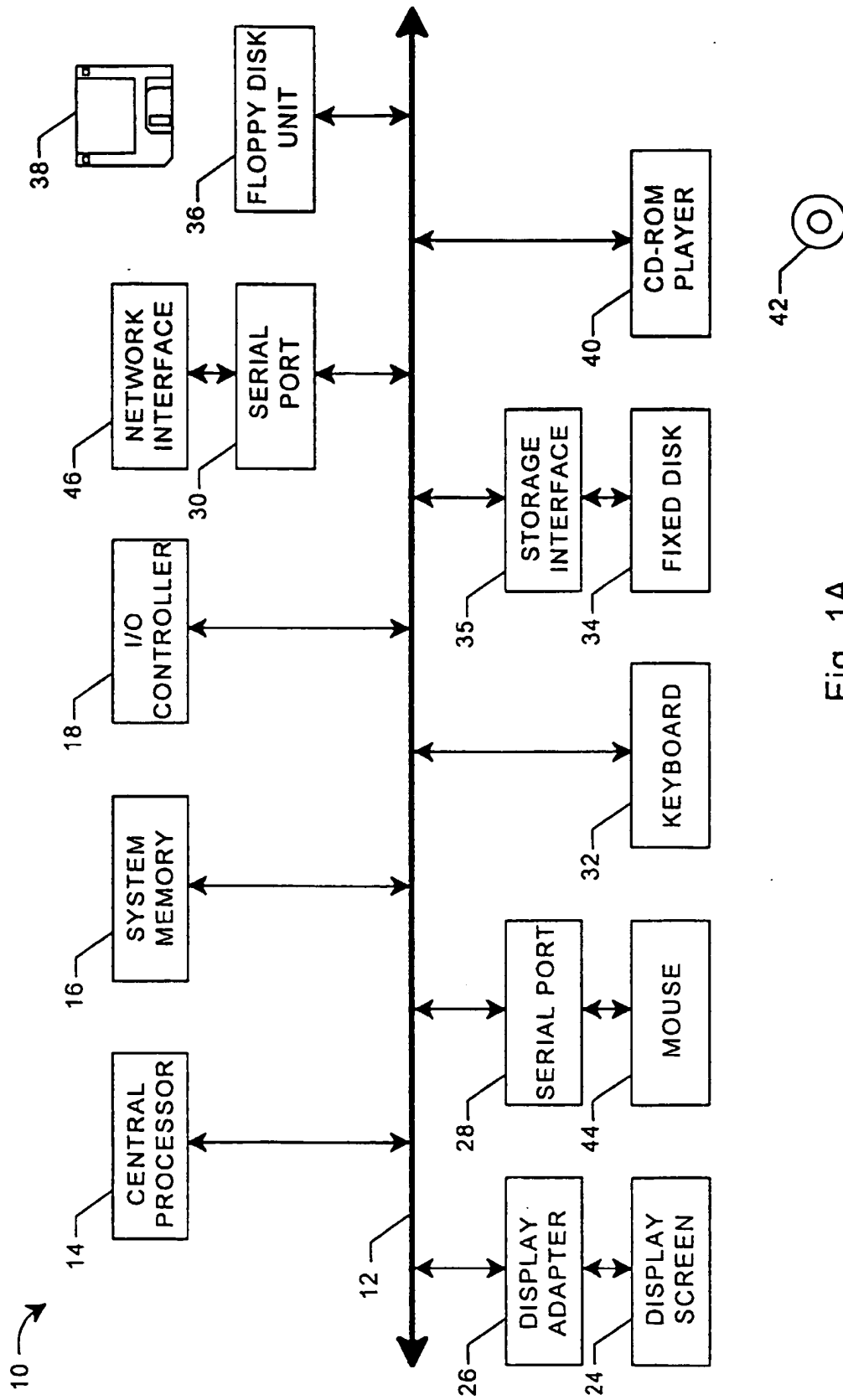


Fig. 1A

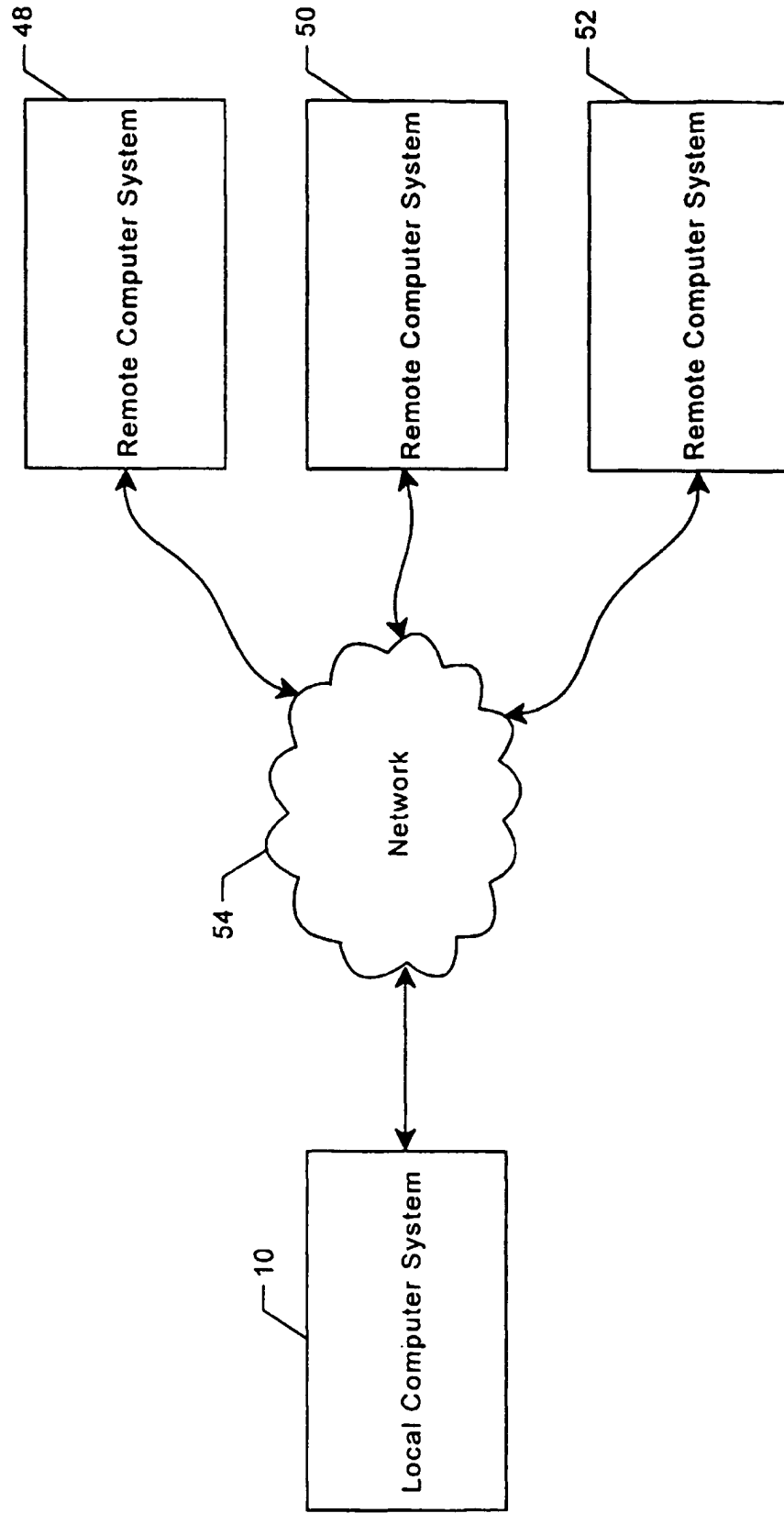


Fig. 1B